

University of Windsor

## Scholarship at UWindor

---

Electronic Theses and Dissertations

Theses, Dissertations, and Major Papers

---

2010

### Depth Map Estimation Using Multi-focus Imaging

Pankajkumar Mendapara  
*University of Windsor*

Follow this and additional works at: <https://scholar.uwindsor.ca/etd>

---

#### Recommended Citation

Mendapara, Pankajkumar, "Depth Map Estimation Using Multi-focus Imaging" (2010). *Electronic Theses and Dissertations*. 131.

<https://scholar.uwindsor.ca/etd/131>

This online database contains the full-text of PhD dissertations and Masters' theses of University of Windsor students from 1954 forward. These documents are made available for personal study and research purposes only, in accordance with the Canadian Copyright Act and the Creative Commons license—CC BY-NC-ND (Attribution, Non-Commercial, No Derivative Works). Under this license, works must always be attributed to the copyright holder (original author), cannot be used for any commercial purposes, and may not be altered. Any other use would require the permission of the copyright holder. Students may inquire about withdrawing their dissertation and/or thesis from this database. For additional inquiries, please contact the repository administrator via email ([scholarship@uwindsor.ca](mailto:scholarship@uwindsor.ca)) or by telephone at 519-253-3000ext. 3208.

# DEPTH MAP ESTIMATION USING MULTI-FOCUS IMAGING

by

Pankajkumar Mendapara

A Thesis

Submitted to the Faculty of Graduate Studies  
through Electrical and Computer Engineering  
in Partial Fulfillment of the Requirements for  
the Degree of Master of Applied Science at the  
University of Windsor

Windsor, Ontario, Canada

2010

© 2010 Pankajkumar Mendapara

# Depth Map Estimation Using Multi-focus Imaging

by

Pankajkumar Mendapara

APPROVED BY:

---

Dr. Dan Wu  
School of Computer Science

---

Dr. Esam Abdel Raheem  
Department of Electrical and Computer Engineering

---

Dr. Q.M. Jonathan Wu  
Department of Electrical and Computer Engineering

---

Dr. Rashid Rashidzadeh  
Chair of Defense

16 April, 2010

# Declaration of Previous Publications

This thesis includes 3 original papers that have been previously published/submitted for publication in peer reviewed conferences, as follows:

Thesis Chapter	Publication title/full citation	Publication status
Chapter 4	Pankajkumar Mendapara, Rashid Minhas, Q. M. Jonathan Wu “Depth Map Estimation Using Exponentially Decaying Focus Measure Based on SUSAN Operator ”, IEEE International Conference on SMC, ©2009 IEEE Reprinted, with permission	Published
Chapter 5	Aryaz Baradarani, Pankajkumar Mendapara, Q. M. Jonathan Wu “Design of a New Class of Odd-Length BiorthogonalWavelet Filter Banks for Signal and Image Processing”, International Conference on Pattern Recognition, ©2010 IEEE Reprinted, with permission	Accepted
Chapter 6	Pankajkumar Mendapara, Aryaz Baradarani, Q. M. Jonathan Wu “Efficient Depth Map Estimation Using Complex Wavelets”, International Workshop on Hot Topics in 3D Multimedia, 2010	Submitted

I certify that I have obtained a written permission from the copyright owner(s) to include the above published material(s) in my thesis. I certify that the above material describes work completed during my registration as graduate student at the University of Windsor.

I declare that, to the best of my knowledge, my thesis does not infringe upon anyones copyright nor violate any proprietary rights and that any ideas, techniques, quotations, or any other material from the work of other people included in my thesis, published or otherwise, are fully acknowledged in accordance with the standard referencing practices. Furthermore, to the extent that I have included copyrighted material that surpasses the bounds of fair dealing within the meaning of the Canada Copyright Act, I certify that I have obtained a written permission from the copyright owner(s) to include such material(s) in my thesis.

I declare that this is a true copy of my thesis, including any final revisions, as approved by my thesis committee and the Graduate Studies office, and that this thesis has not been submitted for a higher degree to any other University or Institution.

# Co-Authorship Declaration

I hereby declare that this thesis incorporates material that is result of joint research, as follows:

This thesis also incorporates the outcome of a joint research undertaken in collaboration with Aryaz Baradarani and Rashid Minhas under the supervision of professor Jonathan Wu. The collaboration is covered in Chapter 4, 5 and 6 of the thesis. In all cases, the key ideas, primary contributions, experimental designs, data analysis and interpretation, were performed by the author, and the contribution of co-author was primarily through the provision of proof reading and reviewing the research papers regarding the technical content.

I am aware of the University of Windsor Senate Policy on Authorship and I certify that I have properly acknowledged the contribution of other researchers to my thesis, and have obtained written permission from each of the co-author(s) to include the above material(s) in my thesis.

I certify that, with the above qualification, this thesis, and the research to which it refers, is the product of my own work.

# Abstract

The research presented in this thesis is concerned with the problem of efficient depth map estimation from multi-focus images acquired at different focus settings. In all, three different depth map estimation techniques are presented. The first method uses SUSAN operator to detect the features, followed by an exponentially decaying function is employed to transfer the distance of the detected features by giving more weight to the nearer vicinity pixels of feature points, which helps to measure the clarity and depth of pixels. A robust, dual-tree complex wavelets and distance transformation based framework is developed for depth map estimation in second focus measure technique. The shift-invariance and better directionality of dual-tree complex wavelets helps to detect the features efficiently, which helps to estimate the depth of the scene more precisely. In third depth map estimation technique, focus measure is ensured by measuring local orientation energy using a quadrature pair of steerable filters of the detected features.

The robustness of proposed depth map estimation techniques are tested by several experiments and results are compared with the other well-documented methods, visually and quantitatively. The obtained results validate the effectiveness of proposed feature based depth map estimation approach.

*to my*  
*wife Suhani and*  
*lovely son Jil*



# Acknowledgements

The Author wishes to acknowledge the guidance given by and patience of his Supervisor, Professor Jonathan Wu during the course of the past two years.

Invaluable help and support provided unselfishly by the members of the Computer Vision and Sensing Systems Laboratory, in particular Aryaz Baradarani, Rashid Minhas, Dibeyndu Mukherjee, Ashirbani Saha, Gaurav Bhatnagar and Abdul Adeel Mohammed gratefully acknowledged.

Author would also like to thank Dr. Esam Abdel Raheem and Dr. Dan Wu, for their valuable suggestions and support during the evaluation of this thesis and seminars.

Author's deepest gratitude goes to his parents, Nagjibhai Mendapara and Nirmalaben Mendapara for their unflagging love and support throughout his life.

Author's family and friends Devendra Mendapara, Praful Mendapara, Bhartesh Dhudashia amongst many others are also gratefully acknowledged for their friendship and understanding in these, often difficult two years.

# Table of Contents

	Page
Declaration of Previous Publications . . . . .	iii
Co-Authorship Declaration . . . . .	v
Abstract . . . . .	vi
Acknowledgements . . . . .	viii
Table of Contents . . . . .	ix
List of Tables . . . . .	xi
List of Figures . . . . .	xii
Abbreviations . . . . .	xiv
<b>1 Introduction . . . . .</b>	<b>1</b>
<b>2 Literature Review . . . . .</b>	<b>5</b>
2.1 Sum of Modified Laplacian Focus Measure . . . . .	5
2.2 Tenenbaum Focus Measure ( $FM_{Tenen}$ ) . . . . .	6
2.3 Gray Level Variance Focus Measure ( $FM_{GLV}$ ) . . . . .	7
2.4 Mean Focus Measure ( $FM_{mean}$ ) . . . . .	7
2.5 Curvature Focus Measure ( $FM_C$ ) . . . . .	8
2.6 Optical Focus Measure ( $FM_O$ ) . . . . .	9
2.7 Principal Component Analysis based Focus Measure ( $FM_{PCA}$ ) . . . .	10
2.8 Approximation Techniques . . . . .	11
<b>3 Fusion Assessment Methods . . . . .</b>	<b>13</b>

3.1	Signal to Noise Ratio, Peak Signal to Noise Ratio and Mean Square Error . . . . .	14
3.2	Mutual Information measure . . . . .	16
3.3	An edge information based objective measure . . . . .	17
3.4	Structural Similarity Index . . . . .	19
<b>4</b>	<b>Depth Map Estimation Using SUSAN Operator and an Exponentially Decaying Function . . . . .</b>	<b>21</b>
4.1	SUSAN Operator . . . . .	22
4.2	Proposed Focus Measure Technique . . . . .	24
4.3	Experiments and Results . . . . .	25
4.4	Conclusion . . . . .	31
<b>5</b>	<b>Depth Map Estimation Using Dual-tree Complex Wavelets and an Exponentially Decaying Function . . . . .</b>	<b>32</b>
5.1	Dual-tree Complex Wavelet Filter Banks . . . . .	32
5.2	Proposed Focus Measure System . . . . .	34
5.3	Experiments and Results . . . . .	36
5.4	Conclusion . . . . .	43
<b>6</b>	<b>Depth Map Estimation Using Dual-tree Complex Wavelets and a Quadrature Pair of Steerable Filters . . . . .</b>	<b>44</b>
6.1	Quadrature Pair of Steerable Filter . . . . .	44
6.2	Proposed Focus Measure Scheme . . . . .	46
6.3	Experiments and Results . . . . .	48
6.4	Conclusion . . . . .	51
<b>7</b>	<b>Conclusion and Future Works . . . . .</b>	<b>52</b>
7.1	Conclusion . . . . .	52
7.2	Future Works . . . . .	53
	<b>References . . . . .</b>	<b>54</b>
	<b>Vita Auctoris . . . . .</b>	<b>57</b>

# List of Tables

4.1	Mutual information comparison . . . . .	30
4.2	Structural similarity comparison . . . . .	30
4.3	Edge information transformation comparison . . . . .	30
5.1	Structural similarity comparison . . . . .	41
5.2	Edge information transformation comparison . . . . .	42
6.1	Structural similarity comparison . . . . .	50
6.2	Edge information transformation comparison . . . . .	50

# List of Figures

1.1	Basic image formation geometry . . . . .	2
1.2	Acquisition setup for an image sequence with varying focus plane . . . . .	3
4.1	Four circular masks at different places on a sample image . . . . .	22
4.2	Four circular masks with similarity coloring; white part of the mask shows the USAN . . . . .	23
4.3	Different steps of proposed algorithm . . . . .	24
4.4	Sample frames of the datasets . . . . .	26
4.5	Depth map of Pre-treatment sequence using various methods . . . . .	27
4.6	Depth map of Chess dataset using various methods . . . . .	28
4.7	Fusion of Chess dataset using various methods . . . . .	29
5.1	A typical two-channel analysis/synthesis dual-tree structure. (a) primal filter bank; (b) dual filter bank. . . . .	33
5.2	Proposed DT-CWT based focus measure system . . . . .	35
5.3	Sample frames of datasets . . . . .	36
5.4	Depth map of Simulated Cone sequence using various methods . . . . .	37
5.5	Depth map of Real Cone dataset using various methods . . . . .	38
5.6	Depth map of Planner Object sequence using various methods . . . . .	38
5.7	Reconstructed 3D shape of Real Cone using various methods . . . . .	39
5.8	Depth map of Real Cone using various methods in presence of Gaussian noise of variance 0.001 . . . . .	40

5.9	Depth map of Planner Object using various methods in presence of shot noise of density 0.001 . . . . .	41
5.10	Comparision of focus measure (Gaussian noise) . . . . .	42
6.1	Block diagram of the proposed technique . . . . .	47
6.2	Depth map of Simulated Cone sequence using various methods . . .	48
6.3	Depth map of Real Cone dataset using various methods . . . . .	49

# Abbreviations

SFF	Shape From Focus
ID	Image Detector
CCD	Charge Coupled Device
FM	Focus Measure
SUSAN	Smallest Univalve Segment Assimilating Nucleus
USAN	Univalve Segment Assimilating Nucleus
$FM_{SML}$	Sum of Modified Laplacian Focus Measure
$FM_T$	Tenenbaum Focus Measure
$FM_{GLV}$	Gray Level Variance Focus Measure
$FM_{mean}$	Mean Focus Measure
$FM_C$	Curvature Focus Measure
$FM_O$	Optical Focus Measure
$FM_{PCA}$	Principal Component Analysis Focus Measure
OTF	Optical Transfer Function
FIS	Focus Image Surface
DP	Dynamic programming
SNR	Signal to Noise Ratio
PSNR	Peak Signal to Noise Ratio
MSE	Mean Square Error
MI	Mutual Information

SSIM	Structural Similarity
EI	Edge Information
MSSIM	Mean Structural Similarity
GT	Geometrical Threshold
DT-CWT	Dual-tree Complex Wavelet Transform
DWT	Discrete Wavelet Transform
SML	Sum of Modified Laplacian
GLV	Gray Level Variance
PCA	Principal Component Analysis



# Chapter 1

## Introduction

The reconstruction of object and to retrieve spatial information from one or multiple observation are crucial issues in computer vision. The automated procedure to extract the structure and depth information of an object or a scene by conveying all the meaningful information from input sensors using mathematical techniques, for human operator or other computer vision tasks, is the main aim of the 3D reconstruction.

When a three dimensional scene is projected onto a two dimensional image plane, depth information is lost. there are two methods to recover this lost depth information, namely, active and passive. In active method using some sensors such as laser or ultrasonic one can directly get the depth information of a scene. These sensors works on the time of flight principal means emits a radiance and measure reflected part. While in passive method such as shape from focus, stereo reconstruction, shape from shading, one can get depth information from multiple input images taken by imaging sensors, using some mathematical techniques.

The shape from focus (SFF) is a technique, utilize to retrieve spatial information from a sequence of images with varying focus plane. In SFF, the depth map is a computation of distance between object points on focus plane and camera lens. The objective of SFF and depth map computation is to reconstruct 3D shape and

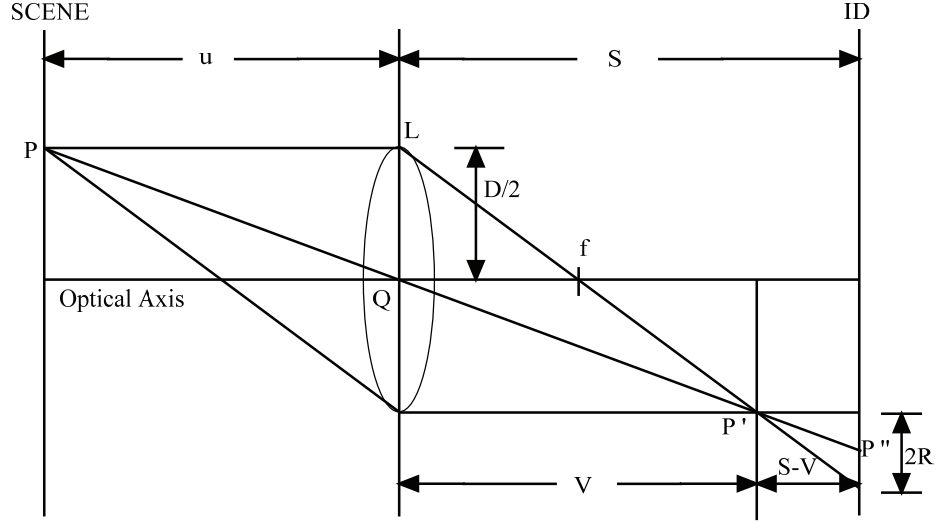


Figure 1.1: Basic image formation geometry

determine depth of every point of an object from the camera lens.

Basic image formation geometry when camera parameters are known is shown in fig. 1.1. As shown in figure, distance of an object from camera lens i.e.  $u$  is required for exact 3D reconstruction of a scene. Depth of a scene, distance of an object from lens, illumination conditions, camera movement, aberration effects in lens and movement in a scene can severely affect the depth map estimation. Computing distance of an object from a camera lens is simple if blur circle radius ( $R$ ) is equal to zero. If image detector  $ID$  is placed at an exact distance  $v$ , sharp focused image  $P'$  is formed. Relationship between object distance  $u$ , focal distance of lens  $f$  and image distance  $v$  is given by Gaussian lens law as,

$$\frac{1}{f} = \frac{1}{u} + \frac{1}{v} \quad (1.1)$$

The object points appears sharp in images which are present on focus plane at the time of acquisition. Focus is an accommodation cue that can be measured from degree of blurring in an image which increases with the distance of imaging

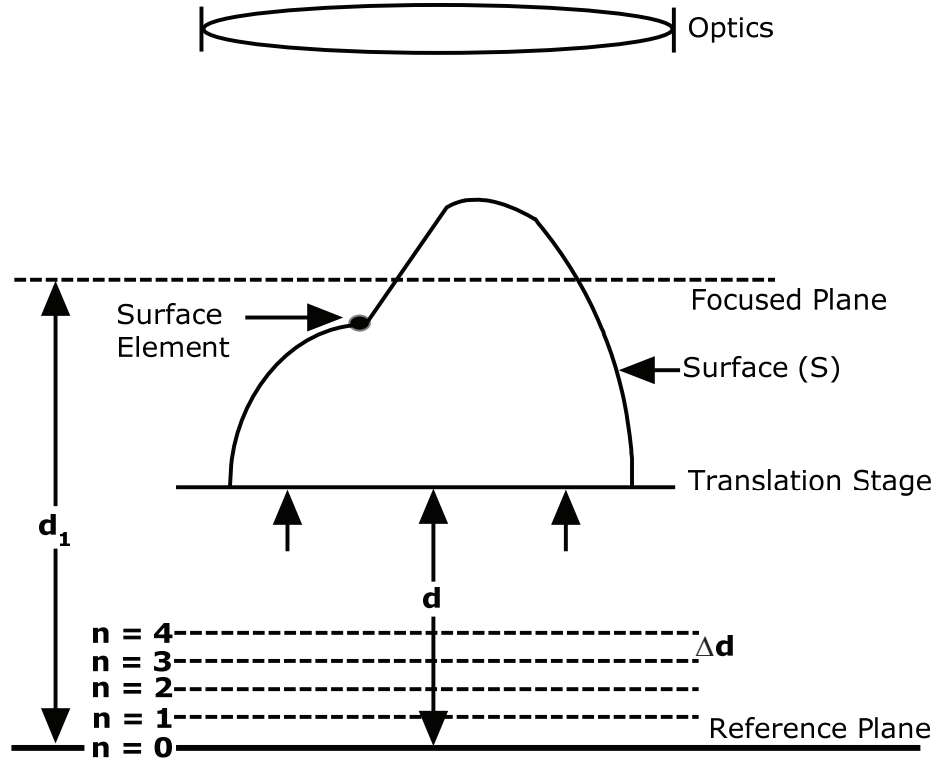


Figure 1.2: Acquisition setup for an image sequence with varying focus plane

system from plane of focus. An image acquisition setup for an image sequence with varying focus plane is shown in fig. 1.2. As shown in figure, object present on a translation stage is imaged at different distance from CCD camera. The object points which are on focus plane have high intensity variation compare to the other points which are not on focus plane. Using any focus measure (FM) technique extract the focused objected points from the image sequence and combining them, gives a composite image. Further, exploiting this information the depth map can be generated. Depth map estimation is widely investigation problem in vision research, has numerous applications such as robot guidance, collision avoidance, 3D feature extraction, medical imaging, seismic data analysis and shape reconstruction.

This thesis contains seven chapters. chapter 2 reviews some well-known focus measure techniques in literature. Some fusion assessment methods which are important to the aspect of evaluation of techniques, are presented in chapter 3. The new methods for depth map estimation, based on SUSAN and exponentially decaying function, based on dual-tree complex wavelets using exponentially decaying functions and based on dual-tree complex wavelets and a quadrature pair of steerable filters are presented in chapter 4, 5 and 6 respectively. Chapter 7 concludes the thesis with an additional insight on future scope of the works.

## Chapter 2

### Literature Review

Multi-focus image fusion has attracted a considerable amount of attention since last two decades. To extract focused points from different images, to construct a composite image and to estimate depth map of a scene are still challenging problems in computer vision. The fundamental concept to estimate the depth map of a scene and to recover the shape by focus analysis is the relationship between focused and defocused images of a scene. A Focus Measure (FM) operator calculates the best focused point in the image, i.e. a FM is defined as a quantity for locally evaluating the sharpness of a pixel. Following we will review some well-known focus measure operators which are proposed in the literature.

#### 2.1 Sum of Modified Laplacian Focus Measure

To measure the quality of focus in a small image area, in [15, 14] a focus measure operator is developed which respond to high frequency variations in image intensity. The Laplacian operator is used for high-pass filtering the image, but the second derivatives in the  $x$  and  $y$  directions can have opposite signs and tends to cancel each other particularly in the case of taxtured image. to overcome this difficulties they defined a modified Laplacian as

$$\nabla^2 I = \frac{\partial^2 I}{\partial x^2} + \frac{\partial^2 I}{\partial y^2} \quad (2.1)$$

$$\nabla_M^2 I = \left| \frac{\partial^2 I}{\partial x^2} \right| + \left| \frac{\partial^2 I}{\partial y^2} \right| \quad (2.2)$$

In order to improve the robustness of weak texture image and for possible variations in the size of texture elements, partial derivative is calculated by variable spacing (step) between the pixels used. The discrete approximation to the modified Laplacian obtained as,

$$\begin{aligned} ML(x, y) &= |2I(x, y) - I(x - step, y) - I(x + step, y)| \\ &+ |2I(x, y) - I(x, y - step) - I(x, y + step)| \end{aligned} \quad (2.3)$$

Finally the focus measure at point  $(i, j)$  is computed as the sum of modified Laplacian values, in a small window around  $(i, j)$ , that are greater than a threshold value

$$FM_{SML}(i, j) = \sum_{x=i-N}^{i+N} \sum_{y=i-N}^{i+N} ML(x, y) \text{ for } ML(x, y) \geq T \quad (2.4)$$

where the parameter  $N$  determines the window size used to compute focus measure,  $I(x, y)$  is the input image frame and  $T$  is a threshold.

## 2.2 Tenenbaum Focus Measure ( $FM_{Tenen}$ )

In [23] a Tenengrade function, to exploits the relationship between a well-focused image and its greater information content, is proposed to measure the sharpness of image pixel. This method is a gradient magnitude maximization method, which measures the sum of the squared responses of the horizontal and vertical Sobel masks. To increase the robustness it summed in a local window as,

$$FM_{Tenen}(i, j) = \sum_{p(x,y) \in U(i,j)} (I_x(x, y)^2 + I_y(x, y)^2)^2 \quad (2.5)$$

where  $I(x,y)$  is the input image frame with the subscript  $x$  and  $y$  denoting the Sobel operator in  $x$  and  $y$  directions respectively, and  $p(x,y)$  is a pixel in the neighborhood  $U(i,j)$  of pixel  $(i,j)$ .

### 2.3 Gray Level Variance Focus Measure ( $FM_{GLV}$ )

The sharp image has higher gray level variance than the blurred image so, the calculation of variance can be serve as a focus measure [9]. The gray level focus measure summed in a local window as,

$$FM_{GLV}(i, j) = \sum_{p(x,y) \in U(i,j)} (I(x, y) - \mu_{U(i,j)})^2 \quad (2.6)$$

where  $I(x,y)$  is the input image frame,  $\mu_{U(i,j)}$  is the mean of gray values in the neighborhood  $U(i, j)$  and the  $p(x,y)$  is a pixel in the neighborhood  $U(i,j)$  of pixel  $(i,j)$ .

### 2.4 Mean Focus Measure ( $FM_{mean}$ )

When the image is becoming sharper the variance of the gray values of that scene is getting higher. The ratio of mean gray value to the center gray value in the neighborhood can be serve to measure the focus pixels in the image [6]. The mean focus measure summed in a local window and can be calculated as,

$$FM(x, y) = \begin{cases} \frac{\mu_{U(i,j)}}{I(x,y)}, & \mu_{U(i,j)} > I(i, j), \\ \frac{I(x,y)}{\mu_{U(i,j)}}, & \text{otherwise} \end{cases} \quad (2.7)$$

$$FM_{mean}(i, j) = \sum_{p(x,y) \in U(i,j)} FM(x, y) \quad (2.8)$$

where  $I(x,y)$  is the input image frame,  $\mu_{U(i,j)}$  is the mean of gray values in the neighborhood  $U(i, j)$  and the  $p(x,y)$  is a pixel in the neighborhood  $U(i,j)$  of pixel  $(i,j)$ .

If the gray value is constant or texture is absent in the image, the ratio is one and if the variation is high the ratio differs from one. If this ratio is between 0 and 1, this factor is inverted.

## 2.5 Curvature Focus Measure (FM<sub>C</sub>)

As stated, the sharper image region implies higher gray value variance than the blurred image region. In [6] the curvature focus measure is proposed by considering the gray values as a 3D surface. The curvature in the sharp image is expected to be higher than the blurred image. to calculate the curvature first the surface is approximated as,

$$f(x, y) = p_0(x) + p_1(y) + p_2(x)^2 + p_3(y)^2 \quad (2.9)$$

where the coefficient  $P = (p_0, p_1, p_2, p_3)^t$  can be found by a least square approximation with  $g_0$  and  $g_2$ .

$$g_0 = \begin{pmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{pmatrix} \quad g_2 = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix} \quad (2.10)$$

$$P = \left( \frac{g_0 * I}{6}; \frac{g_0^t * I}{6}; \frac{3g_2 * I}{10} - \frac{g_2^t * I}{5}; -\frac{g_2 * I}{5} + \frac{3g_2^t * I}{10} \right)^t \quad (2.11)$$



Finally, the curvature focus measure is calculated using the sum of the absolute value of these coefficient as,

$$FM_C(x, y) = |p_0| + |p_1| + |p_2| + |p_3| \quad (2.12)$$

## 2.6 Optical Focus Measure (FM<sub>O</sub>)

In [11], a optical focus measure based on bipolar incoherent image processing is proposed. To measure the sharpness of image pixel the intensity image spectrum is convolved with a optical transfer function. The computed image is given as,

$$I_c(x, y) = Re[|\Gamma_0(x, y)|^2 * h_\Omega(x, y)] \quad (2.13)$$

where, ‘\*’ indicates the convolution,  $|\Gamma_0(x, y)|^2$  is intensity image spectrum and  $h_\Omega(x, y)$  is the transfer function.

The transfer function is basically a optical transfer function (OTF) which can be calculated in frequency domain using fourier transform as,

$$h_\Omega(x, y) = F^{-1}\{OTF_\Omega(k_x, k_y)\} \quad (2.14)$$

where  $k_x$  and  $k_y$  are the spatial frequencies.

The OTF is calculated as,

$$OTF_\Omega(k_x, k_y) = e^{[-\sigma_1(k_x^2 + k_y^2)]} - e^{[-\sigma_2(k_x^2 + k_y^2)]} \quad (2.15)$$

In 2.15  $\sigma_1$  and  $\sigma_2$  are the thresholds (cut-off frequencies), which are calculated as,

$$\sigma_1 = a_1 \left( \frac{f}{k_0} \right)^2 \quad (2.16)$$

$$\sigma_2 = a_2 \left( \frac{f}{k_0} \right)^2 \quad (2.17)$$

where,  $a_1$  and  $a_2$  are constants,  $f$  is the focal length of lens and  $k_0$  is the wavelength of light.

Finally to get the robustness, the optical focus measure is summed in a  $N \times N$  local window as,

$$FM_O(i, j) = \sum_{x=i-N}^{i+N} \sum_{y=i-N}^{i+N} I_C(x, y) \quad (2.18)$$

## 2.7 Principal Component Analysis based Focus Measure ( $FM_{PCA}$ )

In [10, 13], principal component analysis is applied on the multi-focus image sequence to transfer the data into eigenspace and the first feature is employed to calculate the depth value. A vector consisting of 7 neighborhood pixels for each pixel of every frame of image volume is calculated as,

$$V(i, j, k) = [p(i-1, j, k)p(i, j-1, k)p(i, j, k-1)p(i, j, k)p(i+1, j, k)p(i, j+1, k)p(i, j, k+1)]^T \quad (2.19)$$

where,  $i=1,2,\dots,x$ ,  $j=1,2,\dots,y$  and  $k=1,2,\dots,z$  are indices of  $z$  images in the sequence, each of  $X \times Y$  dimensions.

After collecting all such  $z$  vectors, for a pixel  $P(i, j)$  in the  $z$  direction, a matrix of dimensions  $Z \times 7$  is obtained. Let it be denoted by  $M(i, j)$  and its each row contains elements of the vector  $V$ .

PCA is applied on this matrix  $M(i, j)$ . Mean vector  $\mu_n$ , where  $n=1,2,\dots,7$ , and covariance matrix  $C$  are computed as,

$$M(i, j) = [m_{kn}] \quad (2.20)$$

$$\mu_n = \frac{1}{z} \sum_{k=1}^z m_{kn} \quad (2.21)$$

$$C = \frac{1}{z-1} \sum_{k=1}^z (m_{kn} - \mu_n) (m_{kn} - \mu_n)^T \quad (2.22)$$

Eigenvalues and their corresponding eigenvectors are computed for matrix  $C$ . Let  $E$  be the set of eigenvectors then the transformed data  $F$  in eigenspace is obtained by multiplying mean subtracted data with matrix  $E$ .

$$F = E(m_{kn} - \mu_n) \quad (2.23)$$

The columns of the transformed matrix  $F$  are known as principal components or features. It is concluded by analyzing these features that the score of the first feature provides an accurate and robust depth map. The depth value for the pixel  $p(i,j)$  is computed by eq. 2.24. It locates the position of the maximum absolute value in the first component.

$$D(i, j) = Position(max(abs([f_{kl}]))) \quad (2.24)$$

The algorithm runs  $xy$  times to compute the complete depth map.

## 2.8 Approximation Techniques

In depth map estimation because of the discrete number of frames results in some loss of information in between consecutive frames. As a result, the optimum value for some pixels may never be calculated accurately. Hence to address this issue among others, approximation techniques can be applied to the results of the focus measures to construct a more accurate depth map. SFF approximation techniques mentioned in literature, initially use any focus measure and then approximation or learning based techniques is applied to construct more accurate depth map.

In [19, 26] a new concept of focus image surface (FIS) based on planar surface approximations is proposed. They applied this FIS concept on the results of the SML operator. The FIS of an object is defined as the surface formed by the set of

points at which the object points are focused by a camera lens, after first obtaining an estimate of FIS using a traditional SFF method. This estimate is then refined by searching for a planar surface that maximizes the Focus Measure computed over pixels on FIS. To increase the order of surface approximation (quadratic or higher order) and then enlarging the search volume around the initial estimation in [20, 12] a piecewise curved surface approximation of FIS rather than the piecewise planner approximation method is proposed. The piecewise curved surface is estimated by interpolating using a second order Lagrange polynomial. A neural-network is applied on results of GLV operator in [2, 7] to learn the shape of FIS by optimizing the focus measure over small 3D window. In [1] dynamic programming (DP) is applied on the SML results to handle the computational complexity of FIS. This method search for optimal focus measure in whole image volume rather than being limited to a small neighborhood. in [16] 3D shape is estimated using relative defocus blur derived from actual image data to compute the structure of an object.

## Chapter 3

# Fusion Assessment Methods

Since the emergence of image fusion techniques in various applications, methods that can assess or evaluate the performance of different fusion techniques objectively, systematically and quantitatively have been recognized as an important necessity. In this chapter various fusion assessment techniques that have been proposed in the field of image fusion, is discussed.

However, in general, the receiver of the fused images will not be a human viewer but some form of automated image processing system. The loss of some information appears in the fused image compare to the individual source images, and this loss sometimes becomes severe because, this lost information in one particular image processing application might be important for another. Therefore, a general assessment method is always been needed, even if the application of image fusion is unknown in advance.

To assess the fusion performance Li et al. [5] proposed a method to calculate the standard deviation between the reference image (ground truth) and the fused image. Other statistical measurements such as Signal to Noise Ratio (SNR), Peak Signal to Noise Ratio (PSNR) and Mean Square Error (MSE) [21, 28] from digital signal processing are also commonly used to assess the image fusion methods, in case when the reference image is available.

However, in a practical situation the reference image is rarely known. Some image fusion assessment methods are recently developed, which evaluate the fusion method without any reference image. These methods assess the fusion on input-output relationship. In [4] a Mutual Information (MI) principle has been used to assess the fusion method. MI calculates the quantity of information transferred from source images (input) to a fused image. Xydeas and Petrovic [24] Proposed a fusion assessing technique based on pixel level ( $Q_p$ ) in which visual information or perceptual information is directly associated with the edge information while region information is ignored. For the assessment of structural information, a Structural Similarity (SSIM) index Framework is developed in [27] and used to measure the quality of the fused image. These methods (MI, EI and SSIM) measure the amount of information transferred from input images to a fused image and evaluate the fusion method means does not need a reference image.

This chapter is organized as follows. Section 3.1 discusses some statistical measures such as the SNR, PSNR, and MSE, which require an ideal or reference image to assess the fusion technique. A non-linear correlation measures MI and the SSIM index is discussed in section 3.2 and 3.3 respectively. An edge information based objective measure ( $Q_p$ ) is discussed in section 2.4.

### **3.1 Signal to Noise Ratio, Peak Signal to Noise Ratio and Mean Square Error**

Signal to Noise Ratio (SNR), Peak Signal to Noise Ratio (PSNR), and Mean Square Error (MSE) are commonly used measures in assessing image fusion techniques, that consider an image as a special type of signal. The quality of a signal is often expressed quantitatively with the signal to noise ratio defined as [21],

$$\text{SNR} = 10 \log_{10} \left( \frac{E_s}{E_n} \right) \quad (3.1)$$

where  $E_s$  is the sum of the squares of the signal values and  $E_n$  is the sum of the squares of the noise samples. In the context of a signal estimation algorithm, the signal refers to the estimated signal and the noise to the difference (error) between the estimated and the original signal. SNR is unitless and therefore independent of the data units. As far as the image is concerned, the SNR can be written as

$$\text{SNR} = 10 \log_{10} \left( \frac{\sum_{m=1}^M \sum_{n=1}^N z(m, n)^2}{\sum_{m=1}^M \sum_{n=1}^N [z(m, n) - s(m, n)]^2} \right) \quad (3.2)$$

where  $z(m, n)$  and  $s(m, n)$  denote the pixel intensity value of fused and reference image, respectively, at location  $(m, n)$ . The size of the images is  $M \times N$ . High values of SNR show that the error of the estimation is small and, therefore, among various image fusion methods the ones that exhibit higher SNRs can be considered of better performance.

The PSNR and the MSE are measures similar to the SNR and defined as [28],

$$\text{PSNR} = 10 \log_{10} \left( \frac{255^2}{\sum_{m=1}^M \sum_{n=1}^N [z(m, n) - s(m, n)]^2} \right) \quad (3.3)$$

$$\text{MSE} = 10 \log_{10} \left( \frac{\sum_{m=1}^M \sum_{n=1}^N [z(m, n) - s(m, n)]^2}{255^2} \right) \quad (3.4)$$

When assessing the performance of an image fusion technique using the above mentioned measurements, we require knowledge of the original image (ground truth). For that reason these measurements can be used only with synthetic (simulated) data.

The above measurements exhibit the drawback of providing a global idea regarding the quality of an image. In cases where the fused image exhibits artefacts concentrated within a small area, these measurement can still produce an acceptable value even if the image is visually unacceptable.

## 3.2 Mutual Information measure

The measures introduced in the above section are mostly based on the quantitative calculation of the pixel deviation between the original image and the fused image. However, one goal of image fusion is to integrate complementary information from multiple sources so that the fused images are more suitable for the purpose of human visual perception and computer processing. Therefore, a measure should also estimate how much information is obtained from the individual input images. In this section mutual information (MI) measure is presented as a means of assessing image fusion performance.

It is well known that mutual information is a concept from information theory measuring the statistical dependence between two random variables or, in other words, the amount of information that one variable contains about the other. Let  $A$  and  $B$  be two random variables with marginal distributions  $p_A(a)$  and  $p_B(b)$  and joint probability distribution  $p_{AB}(a, b)$ . Mutual information measures the degree of dependence of the two random variables  $A$  and  $B$ . It is defined as follows [4]:

$$\text{MI}_{AB}(a, b) = \sum_{a,b} p_{AB}(a, b) \log \frac{p_{AB}(a, b)}{p_A(a)p_B(b)} \quad (3.5)$$

Considering the image intensity values  $a$  and  $b$  of a pair of corresponding pixels in two images to be samples generated from the random variables  $A$  and  $B$ , respectively, estimation of the joint and marginal distributions  $p_{AB}(a, b)$ ,  $p_A(a)$ , and  $p_B(b)$  can be obtained by normalization of the joint and marginal histograms of both images [4].

The fused image should contain the important information from all of the input (source) images. Obviously the notion of important information depends on the application and is difficult to define. Mutual information is the amount of information that one image contains about another. So, it is possible to employ the mutual information as a measure of image fusion performance. Considering



two input images  $A$ ,  $B$ , and a fused image  $F$ , we can calculate the amount of information that  $F$  contains about  $A$  and  $B$ , according to 3.5

$$\text{MI}_{FA}(f, a) = \sum_{f,a} p_{FA}(f, a) \log \frac{p_{FA}(f, a)}{p_F(f)p_A(a)} \quad (3.6)$$

$$\text{MI}_{FB}(f, b) = \sum_{f,b} p_{FB}(f, b) \log \frac{p_{FB}(f, b)}{p_F(f)p_B(b)} \quad (3.7)$$

Thus, an image fusion performance measure can be defined as

$$\text{MI}_F^{AB} = \text{MI}_{FA}(f, a) + \text{MI}_{FB}(f, B) \quad (3.8)$$

Equation 3.8 indicates that the proposed measure reflects a total amount of mutual information that the fused image  $F$  contains about  $A$  and  $B$ .

### 3.3 An edge information based objective measure

A measure for objectively assessing pixel level fusion performance derived in [24] is presented in this section. The goal in pixel level image fusion is to combine and preserve in a single output image all the 'important' visual information that is present in the input images. Thus, an objective fusion measure ( $Q_p$ ) should extract all the perceptually important information that exists in the input images and measure the ability of the fusion process to transfer as accurately as possible this information into the output image. By evaluating the amount of edge information that is transferred from the individual input images to the fused image, a measure of fusion performance can be obtained. More specifically, consider two input images  $A$  and  $B$  and a resulting fused image  $F$ . A Sobel edge operator is applied to yield the edge strength  $g(m, n)$  and orientation  $a(m, n)$  information for each

image location (pixel)  $(m, n)$ ,  $1 \leq m \leq M$  and  $1 \leq n \leq N$ . Thus, for an input image  $A$

$$g_A(m, n) = \sqrt{S_A^x(m, n)^2 + S_A^y(m, n)^2} \quad (3.9)$$

$$a_A(m, n) = \tan^{-1} \frac{S_A^y(m, n)}{S_A^x(m, n)} \quad (3.10)$$

where  $S_A^x(m, n)$  and  $S_A^y(m, n)$  are the responses of the Sobel masks centred at location  $(m, n)$ .

The relative edge strength and orientation values  $G^{AF}(m, n)$  and  $A^{AF}(m, n)$  of an image  $A$  with respect to an image  $F$  at location  $(m, n)$  are formed as

$$G^{AF}(m, n) = \begin{cases} \frac{g_F(m, n)}{g_A(m, n)}, & g_A(m, n) > g_F(m, n), \\ \frac{g_A(m, n)}{g_F(m, n)}, & \text{otherwise} \end{cases} \quad (3.11)$$

$$A^{AF}(m, n) = 1 - \frac{|a_A(m, n) - a_F(m, n)|}{\pi/2} \quad (3.12)$$

These are used to derive the edge strength and orientation preservation values given below

$$Q_g^{AF}(m, n) = \frac{\Gamma g}{1 + e^{k_g(G^{AF}(m, n) - \sigma_g)}} \quad (3.13)$$

$$Q_a^{AF}(m, n) = \frac{\Gamma a}{1 + e^{k_a(A^{AF}(m, n) - \sigma_a)}} \quad (3.14)$$

where  $Q_g^{AF}g(m, n)$  and  $Q_a^{AF}a(m, n)$  model the perceptual loss of information in  $F$ , in terms of how well the edge strength and orientation values of a pixel  $(m, n)$  in  $A$  are represented in the fused image. The constants  $\Gamma g$ ,  $k_g$ ,  $\sigma_g$  and  $\Gamma a$ ,  $k_a$ ,  $\sigma_a$  determine the exact shape of the sigmoid functions used to form the edge strength and orientation preservation values. Edge information preservation values are then defined as

$$Q^{AF}(m, n) = Q_g^{AF}(m, n) + Q_a^{AF}(m, n) \quad (3.15)$$

with  $0 \leq Q^{AF}(m, n) \leq 1$ . A value of 0 corresponds to complete loss of edge information at location  $(m, n)$  while transferred from  $A$  into  $F$ .  $Q^{AF}(m, n) = 1$  indicates edge information transferred from  $A$  to  $F$  without loss.

Having  $Q^{AF}(m, n)$  and  $Q^{BF}(m, n)$  for  $M \times N$  size images, a normalized weighted performance metric  $Q_P^{AB/F}$  of a given fusion process  $P$  that operates on images  $A$  and  $B$  and produces  $F$  is obtained as follows [24]:

$$Q_P^{AB/F} = \frac{\sum_{m=1}^M \sum_{n=1}^N (Q^{AF}(m, n)\omega^A(m, n) + Q^{BF}(m, n)\omega^B(m, n))}{\sum_{m=1}^M \sum_{n=1}^N (\omega^A(m, n) + \omega^B(m, n))} \quad (3.16)$$

Note that the edge preservation values  $Q^{AF}(m, n)$  and  $Q^{BF}(m, n)$  are weighted by  $\omega^A(m, n)$  and  $\omega^B(m, n)$ , respectively. In general, edge preservation values which correspond to pixels with high edge strength should influence  $Q_P^{AB/F}$  more than those of relatively low edge strength. Thus,  $\omega^A(m, n) = [g_A(m, n)]^L$  and  $\omega^B(m, n) = [g_B(m, n)]^L$ , where  $L$  is a constant. Also note that  $0 \leq Q_P^{AF}(m, n) \leq 1$ .

### 3.4 Structural Similarity Index

Generally, natural image signals carry important information about the structure of object in visual scene. In [27] a framework, to measure the distortion of the fused image from the source images, so called, structural similarity (SSIM) index is proposed. The SSIM index measures how much the structure of the fused image is similar to that of the input images.

To measure the similarity of the structure this framework uses three elements of image patches; the similarity of brightness, contrast and structure. The similarity calculation of brightness, contrast and structure is calculated by using eq. 3.17, eq. 3.18 and eq. 3.19 respectively as,

$$l(h_i, z) = \frac{2\mu_{h_i}\mu_z + C_1}{\mu_{h_i}^2 + \mu_z^2 + C_1} \quad (3.17)$$

$$c(h_i, z) = \frac{2\sigma_{h_i}\sigma_z + C_2}{\sigma_{h_i}^2 + \sigma_z^2 + C_2} \quad (3.18)$$

$$s(h_i, z) = \frac{\sigma_{h_i z} + C_3}{\sigma_{h_i}\sigma_z + C_3} \quad (3.19)$$

where  $h_i$ ,  $i \in 1, 2, \dots, N$  represents input image sequence and  $z$  is the fused image while,  $\mu$  and  $\sigma$  are mean intensity and standard deviation respectively.  $\sigma_{h_i z}$  is the sample cross correlation of  $h_i$  and  $z$  after removing their mean.  $C_1$ ,  $C_2$  and  $C_3$  are small positive constants that stabilize each term so that near sample means, variance or correlation does not lead to numerical instability.

Finally, combining these three elements of image patches, the SSIM and mean SSIM index forms as,

$$SSIM(h_i, z) = \frac{(2\mu_{h_i}\mu_z + C_1)(2\sigma_{h_i z} + C_2)}{(\mu_{h_i}^2 + \mu_z^2 + C_1)(\sigma_{h_i}^2 + \sigma_z^2 + C_2)} \quad (3.20)$$

$$MSSIM = \sum_{h_i=1}^N \text{mean2}(SSIM(h_i, z)) \quad (3.21)$$

where  $N$  is the number of input images.

## Chapter 4

# Depth Map Estimation Using SUSAN Operator and an Exponentially Decaying Function

In an image processing, detection of features have significance importance. In multi-focus image fusion accurate detection of features plays a pivotal role. More accurately detected features enable the fusion perfectly from different multi-focus input images and concurrently help to calculate the depth of a scene or an object. The absolute depth of object surface patches can be calculated from the focal length and the position of lens that gave the sharpest image of the surface patches.

A novel depth map estimation technique using an exponentially decaying function based on SUSAN operator is proposed in this chapter. The SUSAN, a non-linear low-level image processing algorithm is used for feature detection in the proposed technique. The excellent capacity of edge detection, corner detection and structure preserving noise filtering of SUSAN, helps to identify features in multi-focus image efficiently and exponentially decaying function helps to measure focus by changing pixel value according with the distance from feature pixel, gives more weight in the nearer vicinity of the feature pixel.

## 4.1 SUSAN Operator

Smith and Brady proposed SUSAN (Smallest Univalve Segment Assimilating Nucleus) algorithm in 1997 [18]. This algorithm has three parts: edge detection, corner detection and structure preserving noise filtering. In this algorithm non-linear filtering is used to identify image sub-regions which are closely related to individual pixels. A circular mask is used for convolution and the brightness of each pixel within the circle is compared with the brightness of center pixel (nucleus) of the mask. The area of the mask that has the same brightness as the nucleus is known as USAN (Univalve Segment Assimilating Nucleus).

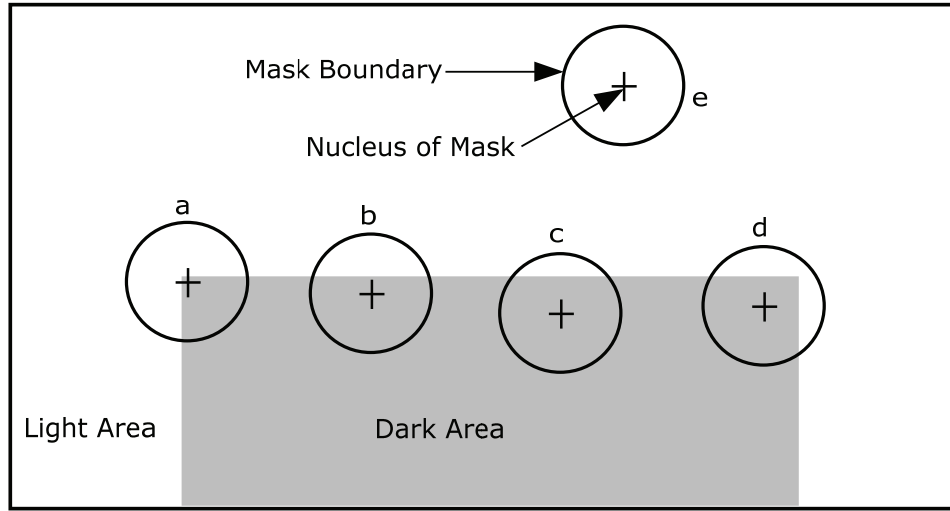


Figure 4.1: Four circular masks at different places on a sample image

In fig. 4.1 a circular mask is placed at four different places in an image. The USAN area is at a maximum when nucleus lies within the flat region as shown at place (c), it falls half of this maximum when nucleus is nearer to edge (at place (b)) and falls further nearer to corner as shown in fig. 4.1 at place (a) and (d). The detected USAN area is shown in fig. 4.2. The area of USAN conveys the most important information about the structure of the image. Thus, by knowing the USAN area i.e. from size, centroid and second moment of the USAN, edges and

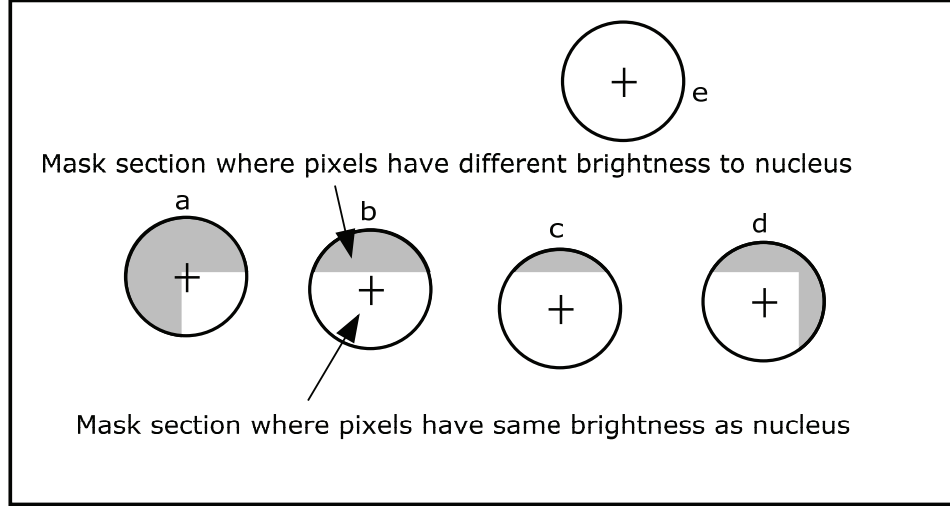


Figure 4.2:  
Four circular masks with similarity coloring; white part of the mask shows the USAN

two dimensional features (corners) can be easily detected.

Let,  $I(r)$  denotes the gray value at pixel  $r$ ,  $n$  the area of the USAN (the total no. of pixels in the USAN),  $r_0$  the nucleus,  $\alpha$  is circular mask and  $\sigma$  is brightness difference threshold, then

$$n(r_0) = \sum_{r \in \alpha} c(r, r_0) \quad (4.1)$$

Finally, the response of SUSAN operator at pixel  $r_0$  is given by,

$$E(r_0) = \begin{cases} GT - n(r_0) & \text{for } n(r_0) < GT, \\ 0 & \text{else} \end{cases} \quad (4.2)$$

where,  $GT$  is called geometrical threshold. For edge detection suitable value of  $GT$  is  $(\frac{3}{4}) n_{max}$  where  $n_{max}$  is the maximum value that  $n$  can carry.

## 4.2 Proposed Focus Measure Technique

The clarity measurement of focused pixels of the multi-focus image sequence, is a vital importance for any focus measure technique. Successiveness of any focus measure technique depends on its ability to calculate the sharpness value of each image pixel. Proposed scheme uses an exponentially decaying function with SUSAN operator to analyze the sharpness of each pixel in an image sequence. Entire process of proposed scheme is depicted in fig. 4.3.

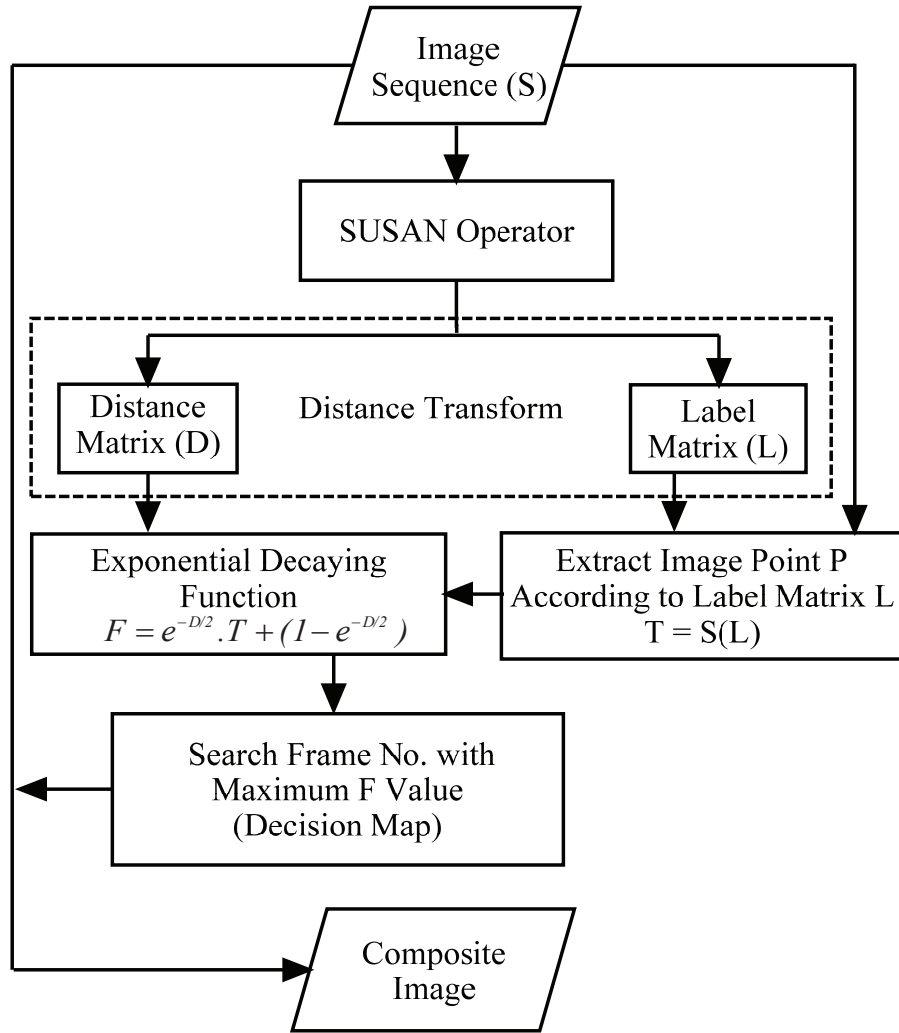


Figure 4.3: Different steps of proposed algorithm



In proposed technique, features are extracted by applying SUSAN operator on the pre-registered multi-focus image sequence. A focus measure characterized by exponentially decaying function, is employed to compute sharpness of each pixel in an image. such decaying function uses neighborhood information of extracted feature points assuming that the pixel value far from a feature points is equal to 1 and it approaches this limiting value in an exponential way.

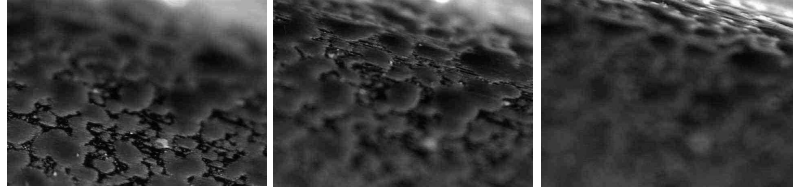
Given an arbitrary point  $k$  and the set  $E$  of feature points, the focus measure  $F$  is estimated as,

$$F(x, y) = e^{-\frac{D(x, y)}{2}} T(x, y) + \left(1 - e^{-\frac{D(x, y)}{2}}\right) \quad (4.3)$$

where,  $D$  is the distance between point  $k$  and the nearest feature pixel  $T$ . Distance transformation of an edge image which is an output of SUSAN operator yields a distance matrix ( $D$ ) and a label matrix ( $L$ ).  $T$  is the actual intensity value extracted from original images using label matrix  $L$ . These matrices  $D$  and  $T$  are used in computation of  $F$  values for each frame.(eq. 4.3). After changing pixel value using an exponentially decaying function, the decision map is computed by comparing pixels in  $Z$  direction. The frame with higher value of  $F$ , that frame number is mapped on to corresponding pixels of the decision map. Finally using decision map, pixels are extracted from original image sequence yielding a composite image.

### 4.3 Experiments and Results

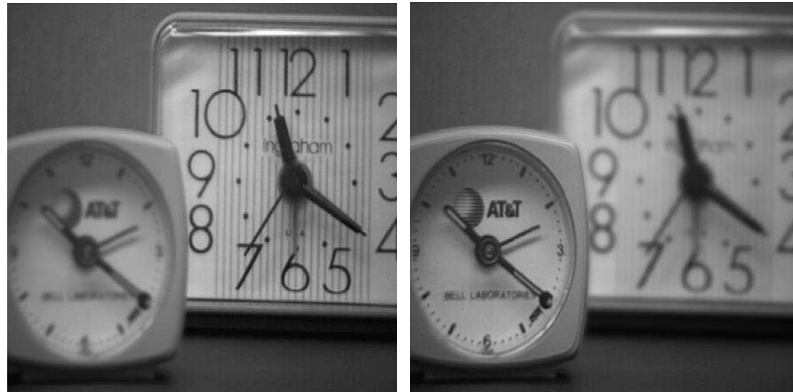
The performance of proposed depth map estimation technique is tested with various multi-focus pre-registered image datasets, namely, pre-treatment, chess, clock and book, visually as well as quantitatively. For quantitative measurement mutual information (MI), structural similarity (SSIM) index and edge information transformation matrix (Qp) are used. These methods assess the fusion on the basis of



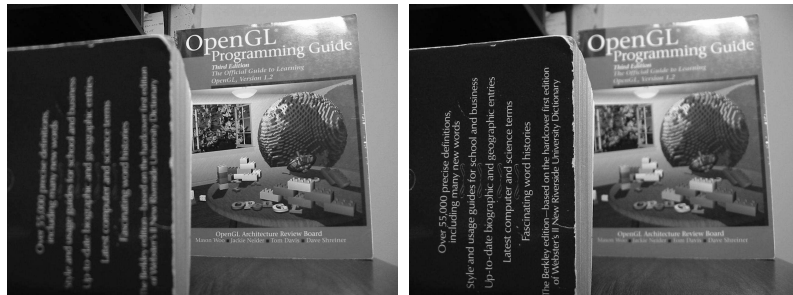
(a) Pre-treatment



(b) Chess



(c) Clock



(d) Book

Figure 4.4: Sample frames of the datasets

input-output relationship means does not need the reference image.

The pre-treatment data set have 8 partially focused images. which is taken by moving camera lens stepwise diagonally, with image resolution  $800 \times 600$ , the chess image sequence is obtained by moving lens horizontally in 29 steps, and each image  $800 \times 600$  size, clock and book datasets have two images, focus different parts in each frame. The image size of clock and book are  $512 \times 512$  and  $1280 \times 960$  respectively. Sample frames of all datasets used in experiment is shown in fig 4.4.

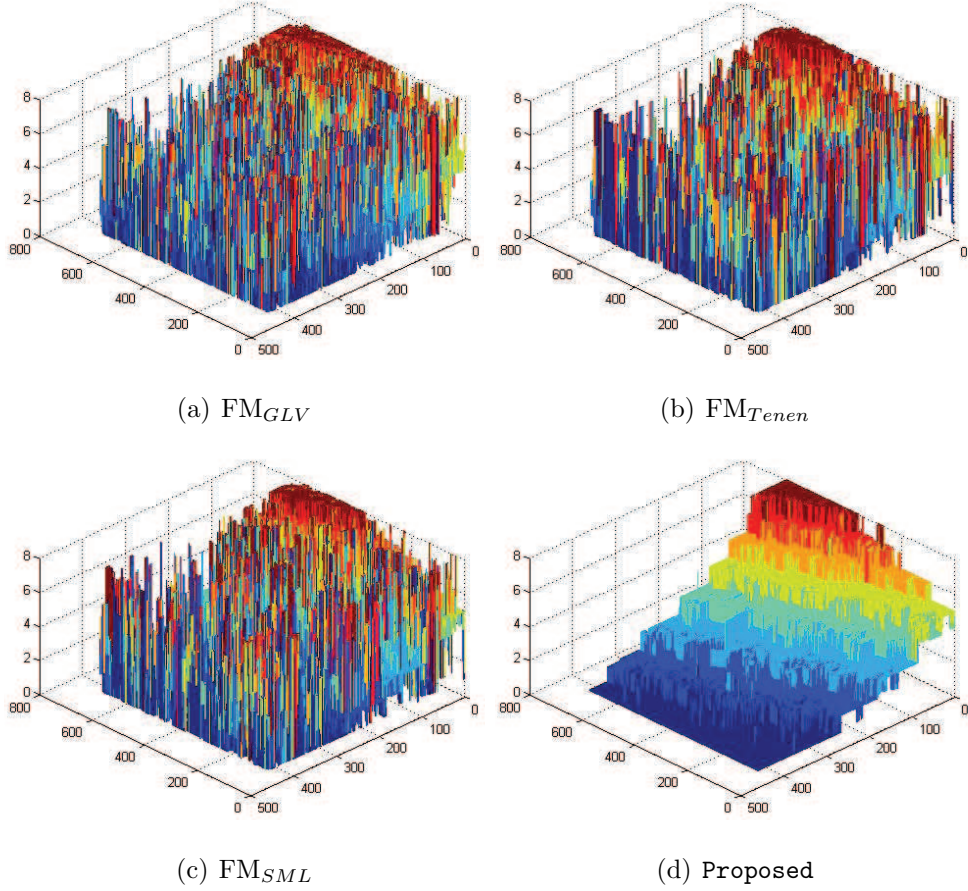


Figure 4.5: Depth map of Pre-treatment sequence using various methods

The depth map obtained for pre-treatment sequence and chess dataset using proposed focus measure system and other three well-known  $FM_{GLV}$ ,  $FM_{Tenen}$ ,  $FM_{SML}$  are shown in fig. 4.5 and fig. 4.6 respectively. As shown in fig. 4.5, the

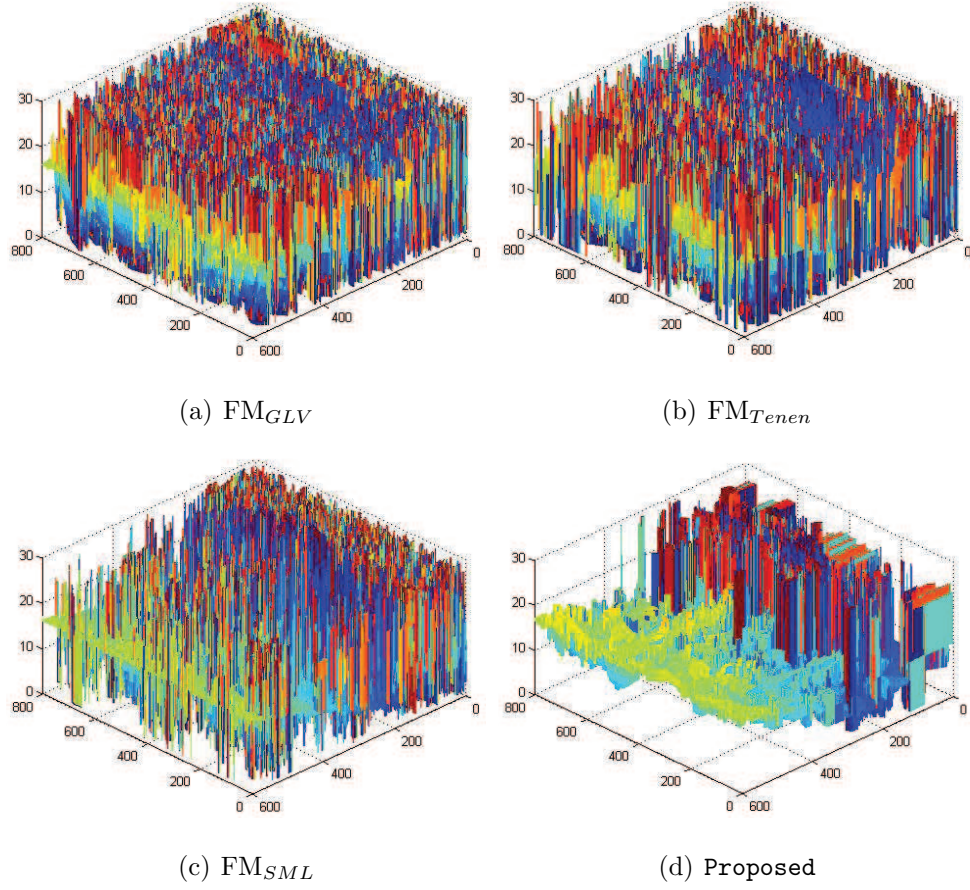
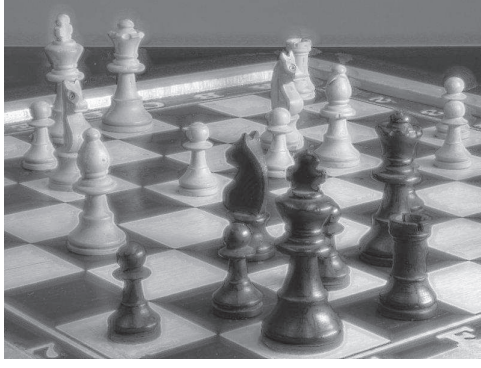


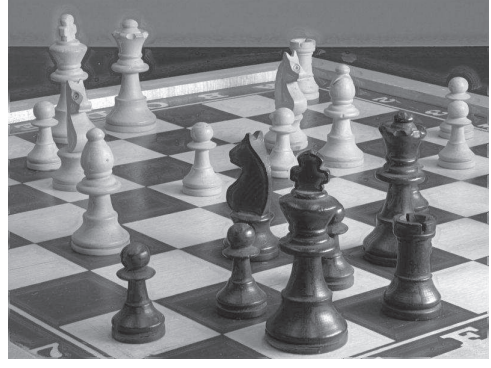
Figure 4.6: Depth map of Chess dataset using various methods

obtained depth map using proposed method is smoother than the other three well-documented methods for pre-treatment dataset. The steps of each movement of the camera lens is clearly visible in the depth map of proposed method. Similarly, for chess dataset the performance of the proposed method is quite better and estimate the depth of the scene more accurately compare to other methods as shown in fig. 4.6.

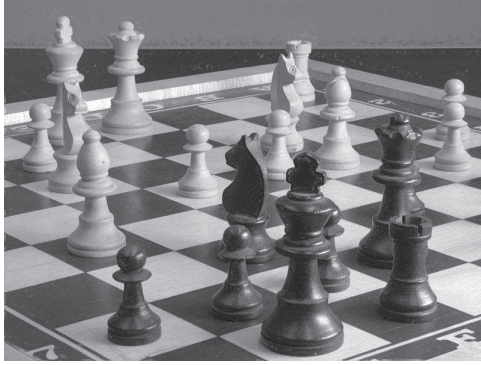
Fig. 4.7 shows the final fused image of the chess image sequence using proposed SUSAN operator based focus measure technique and three other methods. From figure it is seen that the fusion obtained by proposed method have more brightness-contrast ratio than the other methods as well as at edges, performance is quite better without any hall effect, while other methods simply fail.



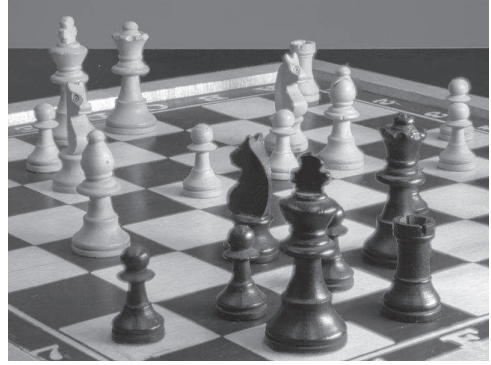
(a)  $FM_{GLV}$



(b)  $FM_{Tenen}$



(c)  $FM_{SML}$



(d) Proposed

Figure 4.7: Fusion of Chess dataset using various methods

For quantitative evaluation of the proposed method we have used three different criteria; mutual information (MI), Structural Similarity (SSIM) index and edge information transformation matrices ( $Q_p$ ). The obtained result in terms of MI, MSSIM and  $Q_p$  of the proposed and three other  $FM_{GLV}$ ,  $FM_{tenen}$ ,  $FM_{SML}$  methods for pre-treatment, chess, clock and book dataset without noise and in presence of Gaussian noise of variance 0.0005 are shown in table 4.1, table 4.2 and table 4.3 respectively.

From table 4.1, 4.2 and 4.3; it is evident that the performance of proposed method in terms of mutual information, structural similarity and edge transformation is better than the other well-documented methods.

	$FM_{GLV}$	$FM_{Tenen}$	$FM_{SML}$	Proposed
Pre-treat	2.6076	2.7035	2.6887	2.9439
Chess	13.3120	13.3002	15.3308	17.4559
Clock	2.3051	2.3187	2.3499	2.4215
Book	2.3865	2.3976	2.3975	2.4235

Table 4.1: Mutual information comparison

	$FM_{GLV}$	$FM_{Tenen}$	$FM_{SML}$	Proposed
Pre-treat	0.5414	0.5448	0.5455	0.5735
Chess	0.7111	0.6757	0.7082	0.7717
Clock	0.8630	0.8702	0.8797	0.8929
Book	0.8662	0.8675	0.8728	0.8840

Table 4.2: Structural similarity comparison

	$FM_{GLV}$	$FM_{Tenen}$	$FM_{SML}$	Proposed
Pre-treat	0.3473	0.3944	0.3875	0.4034
Chess	0.3172	0.2421	0.2449	0.3184
Clock	0.7033	0.7554	0.7667	0.7964
Book	0.6788	0.7466	0.7487	0.7656

Table 4.3: Edge information transformation comparison

## 4.4 Conclusion

In this chapter, a new focus measure method for depth map estimation based on exponentially decaying function, that exploits neighborhood information of extracted feature points identified through SUSAN operator is presented. Structure preserving noise filtering and detection of various kinds of features (edges, corners) using SUSAN allows improved detection of well focused image points while exponentially decaying function gives more weight to the pixels lie in the nearer vicinity of the feature pixels. Experimental results show the superior performance of proposed method compare with other traditional scheme. Medical imaging, collision avoidance, shape reconstruction and object segmentation are some of the potential areas that can benefit from scheme presented in this chapter.



## Chapter 5

# Depth Map Estimation Using Dual-tree Complex Wavelets and an Exponentially Decaying Function

Fundamental to the concept of depth estimation by focus analysis is the relationship between focused and defocused images of a scene. The depth of every point of an object from the camera lens is calculated by finding the best focused points. In this chapter an efficient depth map estimation technique based on DT-CWT and an exponentially decaying function is proposed. Shift-invariance nature and better directionality of DT-CWT helps to detect the oriented features very efficiently while distance transformation using exponentially decaying function gives more weight to the neighborhood pixel of the feature points.

### 5.1 Dual-tree Complex Wavelet Filter Banks

Dual-tree complex wavelet transform (DT-CWT) which is an enhancement to the discrete wavelet transform (DWT), possesses two key properties, i.e., the transformation is nearly shift invariant and it has better directionality in higher-dimensional space. As a typical DWT structure is a well known transformation,



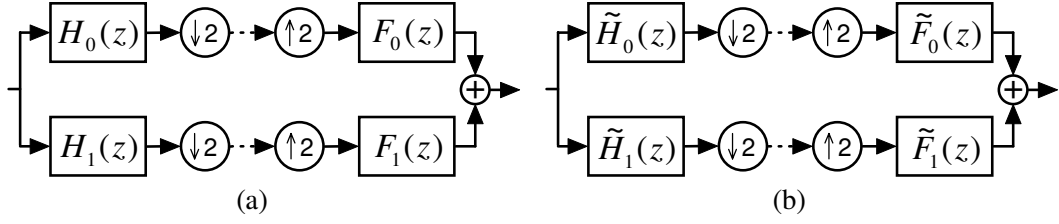


Figure 5.1:

A typical two-channel analysis/synthesis dual-tree structure. (a) primal filter bank; (b) dual filter bank.

in this section, a brief review is given for DT-CWT only. Consider a two-channel dual-tree filter bank implementation of the complex wavelet transform. Shown in Figure 5.1(a), the primal filter bank  $\mathbf{B}$  in each level defines the real part of the wavelet transform. The dual filter bank  $\tilde{\mathbf{B}}$  shown in Figure 5.1(b) defines the imaginary part. Note that the scaling and wavelet functions associated with the analysis side of  $\mathbf{B}$  are defined by the iterative two-scale equations

$$\phi_h(t) = 2 \sum_n h_0[n] \phi_h(2t - n) \quad (5.1)$$

$$\psi_h(t) = 2 \sum_n h_1[n] \phi_h(2t - n) \quad (5.2)$$

The scaling function  $\phi_f$  and wavelet  $\psi_f$  in the synthesis side of  $\mathbf{B}$  are similarly defined with  $f_0$  and  $f_1$  and the same discussion is valid for the scaling functions ( $\tilde{\phi}_h$  and  $\tilde{\phi}_f$ ) and wavelet functions ( $\tilde{\psi}_h$  and  $\tilde{\psi}_f$ ) of the dual filter bank  $\tilde{\mathbf{B}}$ . The dual-tree filter bank defines analytic complex wavelets  $\psi_h + j\tilde{\psi}_h$  and  $\tilde{\psi}_f + j\psi_f$ , if the wavelet functions of the two filter banks form Hilbert transform pairs. Specifically, the analysis wavelet  $\tilde{\psi}_h(t)$  of  $\tilde{\mathbf{B}}$  is the Hilbert transform of the analysis wavelet  $\psi_h(t)$  of  $\mathbf{B}$ , and the synthesis wavelet  $\psi_f(t)$  of  $\mathbf{B}$  is the Hilbert transform of  $\tilde{\psi}_f(t)$ . In other words,

$$\tilde{\Psi}_h(\omega) = -j \text{sign}(\omega) \Psi_h(\omega) \quad (5.3)$$

$$\Psi_f(\omega) = -j \text{sign}(\omega) \tilde{\Psi}_f(\omega) \quad (5.4)$$

where  $\Psi_h(\omega)$ ,  $\Psi_f(\omega)$ ,  $\tilde{\Psi}_h(\omega)$ , and  $\tilde{\Psi}_f(\omega)$  are the Fourier transforms of wavelet functions  $\psi_h(t)$ ,  $\psi_f(t)$ ,  $\tilde{\psi}_h(t)$ , and  $\tilde{\psi}_f(t)$  respectively, and **sign** represents the signum function.

There are two traditional approaches to the design of Hilbert pairs of complex wavelet bases. One is to design both the primal and the dual banks simultaneously, e.g., Kingsbury's q-shift [8] and Selesnick's common factor [17] solutions. The other method, known as matching technique, is to design the dual filter bank for an existing filter bank [22]. In general, the transformation is approximately shift-invariant and provides directionality-selective subbands in wavelet domain [8, 17] while preserving perfect reconstruction. It is worth pointing out that these properties are missing in discrete wavelet transform.

## 5.2 Proposed Focus Measure System

In general, ability of any focus measure technique to calculate the sharpness value of each image pixel shows the success of a method. To calculate the sharpness value of pixels and subsequently to find the depth of each pixel, we propose the use of dual-tree complex wavelet transform along with the so-called distance transformation employing an exponentially decaying function. Because of better directionality and shift-invariance the DT-CWT helps to detect the feature information accurately and the distance transformation employing exponentially decaying function helps to give more weight to the pixel lies in the near vicinity of the feature pixels. The entire block diagram of the proposed DT-CWT based focus measure system is depicted in fig. 5.2.

As is seen in this figure, the entire multi-focus image sequence is decomposed via DT-CWT in to scaling and wavelet subbands using the 9/7–10/8 filter banks proposed in [25]. The decomposition using DT-CWT provides detailed information for each frame in the sequence; leading to more options to select the best

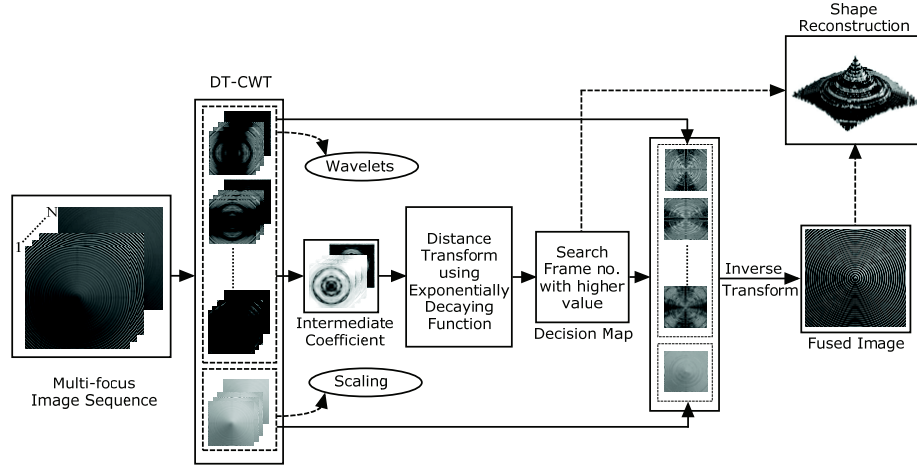


Figure 5.2: Proposed DT-CWT based focus measure system

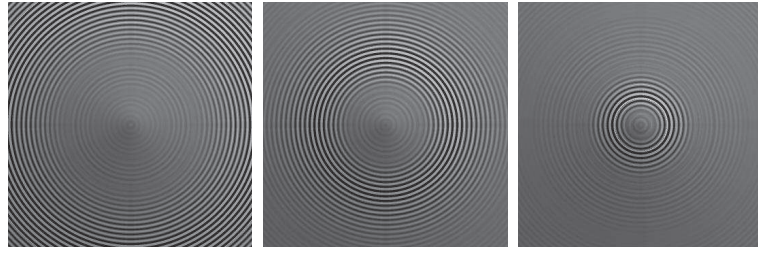
fitting feature among the entire input images. It is technically an important task to know which pixel comes from which image. Also, in order to preserve information and to avoid pixels indices complexity, we do not perform downsampling between stages while decomposing data by DT-CWT and thus each subband has the same size as an input image. After decomposing multi-focus image sequence with DT-CWT the intermediate subband ( $I$ ) with higher pixel value is computed applying maximum selection rule on subbands of each image in wavelet domain. Finally, a decision map is determined employing the distance transform with the exponentially decaying function as,

$$F(x, y) = e^{-\frac{d(x, y)}{2}} I(x, y) + \left(1 - e^{-\frac{d(x, y)}{2}}\right) \quad (5.5)$$

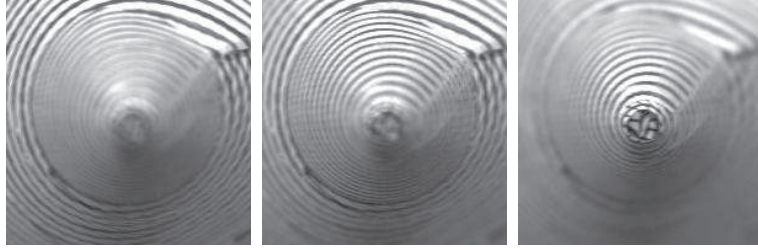
The exponentially decaying function uses neighborhood information in each subband assuming that the pixel value far from a pixel is equal to 1 and it approaches this limiting value exponentially. The decision map for fusion is constructed by comparing the value of subbands; the frame number with higher value is mapped onto the decision map. Using the decision map, pixels are extracted from corresponding subbands of the image sequence and an inverse DT-CWT yields the final fused image.

### 5.3 Experiments and Results

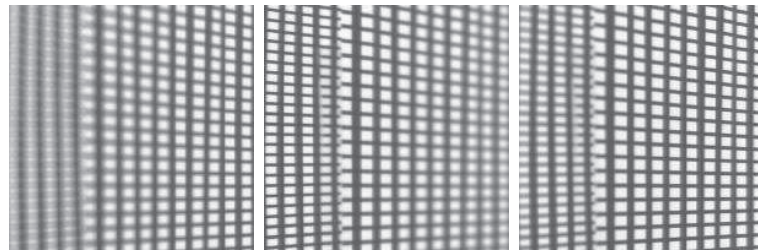
In experiments three different kind of multi-focus image sequences, namely, simulated cone, real cone and planner object are used. The simulated cone sequence is generated using a software while the real cone and planner object sequence is taken by a CCD camera. The simulated cone and real cone sequences have 97 partially focused images with image size  $360 \times 360$  and  $200 \times 200$ , while the planner object sequence has 87 multi-focused  $200 \times 200$  size images. The sample frames of all three datasets are shown in fig. 5.3.



(a) Simulated Cone



(b) Real Cone



(c) Planner Object

Figure 5.3: Sample frames of datasets

The effectiveness of the proposed DT-CWT based focus measure system is tested by comparing with three other well-known focus measure methods, namely, tenengrade ( $FM_{Tenen}$ ), gray level variance ( $FM_{GLV}$ ) and sum of modified laplacian ( $FM_{SML}$ ) proposed in literature. The resultant depth map obtained by proposed and three other methods for simulated cone, real cone and planner object are shown in fig. 5.4, fig. 5.5 and fig. 5.6 respectively. From these figures it is clear that the depth maps of the proposed method are more smoother and uniform compare to the other methods. Note that the results obtained using DT-CWT are free of spikes and the boundaries of depth are clearly separable, which indicates that proposed method is more accurate to judge the focused pixel from the partially focused images.

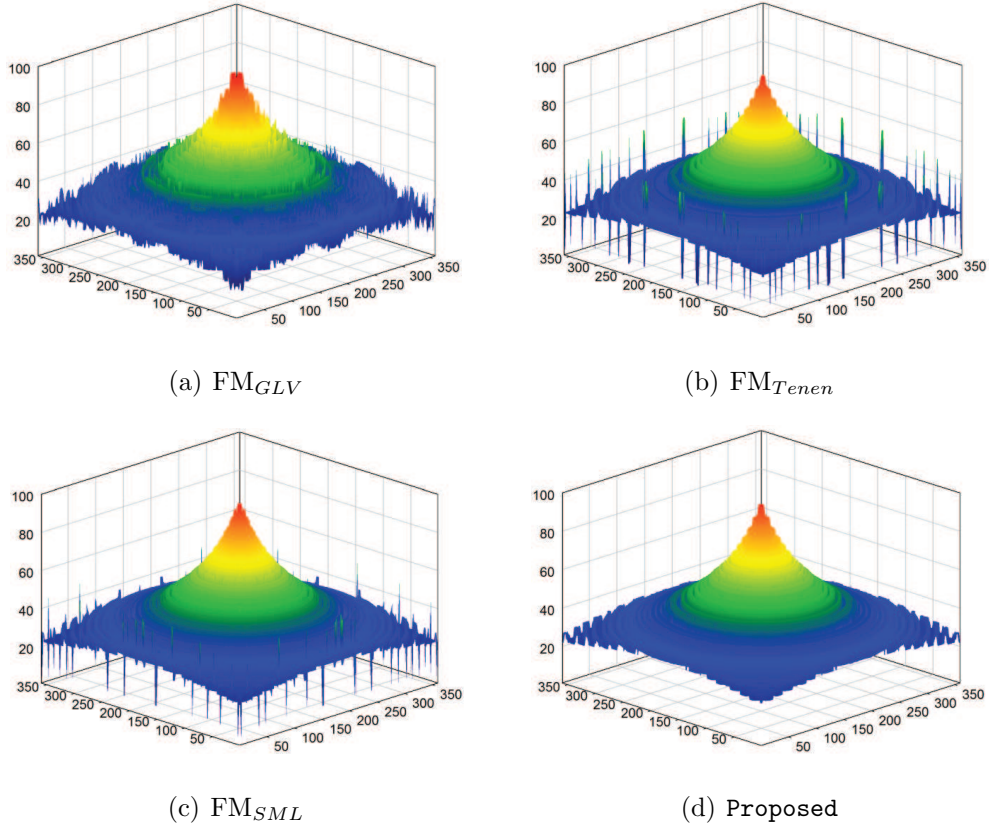


Figure 5.4: Depth map of Simulated Cone sequence using various methods

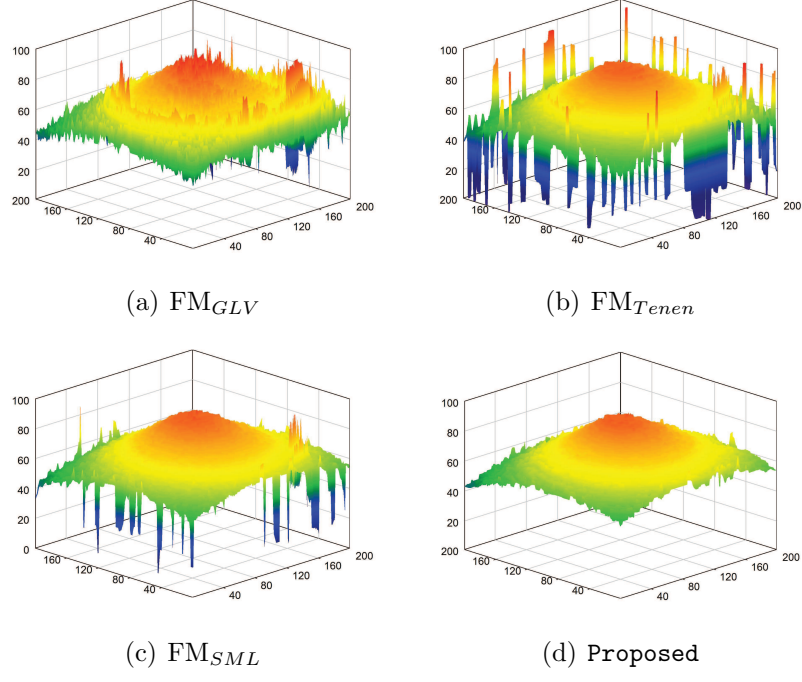


Figure 5.5: Depth map of Real Cone dataset using various methods

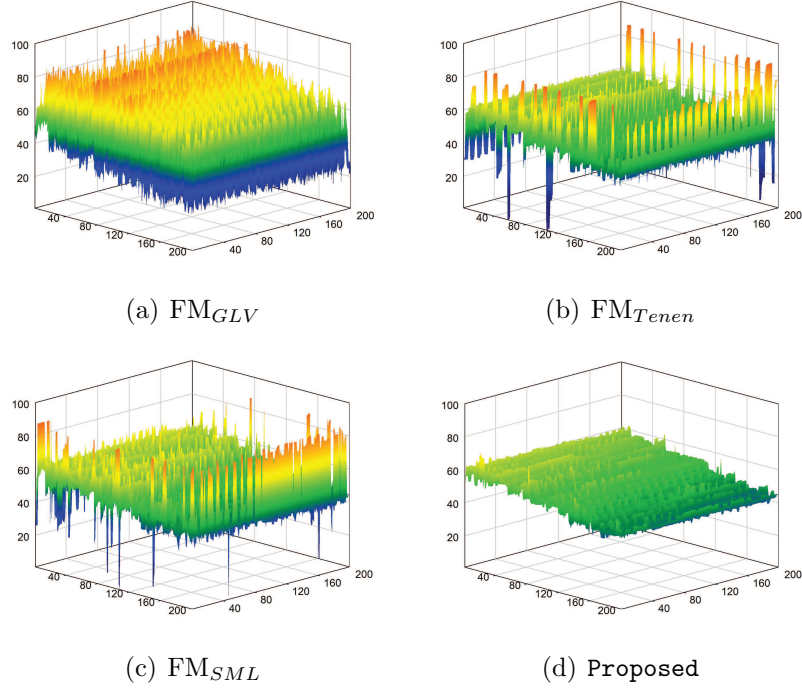


Figure 5.6: Depth map of Planner Object sequence using various methods

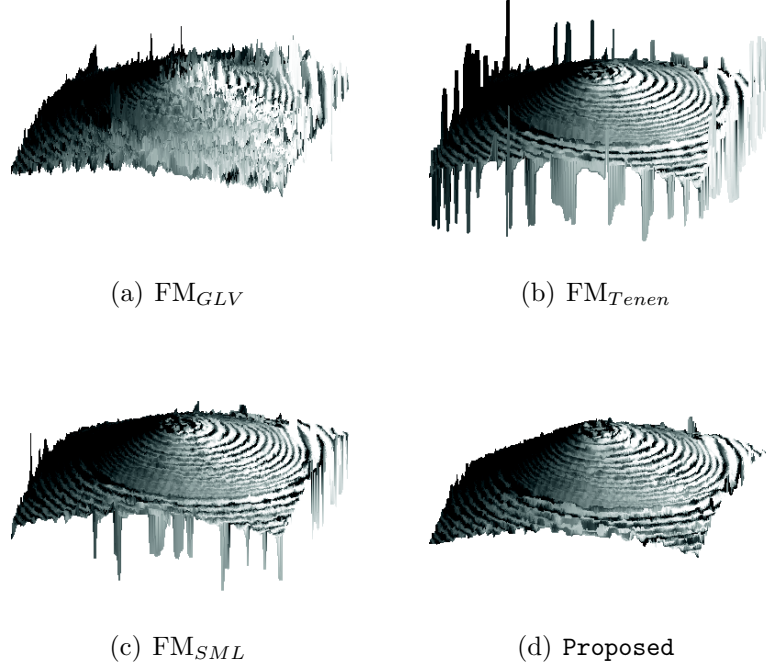


Figure 5.7: Reconstructed 3D shape of Real Cone using various methods

Once, the fused image and depth map are available then the 3D shape reconstruction is straight. The reconstructed 3D shape of real cone using various techniques are shown in fig. 5.7

In image processing noise have significance importance. Some focus measure system, namely,  $FM_{Tenen}$ ,  $FM_{SML}$  and  $FM_{GLV}$  uses derivatives to calculate sharpness value of the pixel. Derivation is very much sensitive to noise, resulting drastic degradation of the performance of the focus measure system.

Thermal noise is the main contributor to CCD noise, which usually introduce during image acquisition can be modeled using Gaussian PDF. So, to check the robustness of the proposed focus measure system, a Gaussian noise of variance 0.001 is added to the real cone dataset, resultant depth map obtained using various methods is depicted in fig. 5.8. From figure it is seen that in presence of noise  $FM_{Tenen}$  and  $FM_{GLV}$  methods exhibits poor performance and the depth map of the  $FM_{SML}$  is even unrecognizable, while the proposed method still track the actual depth of the scene.



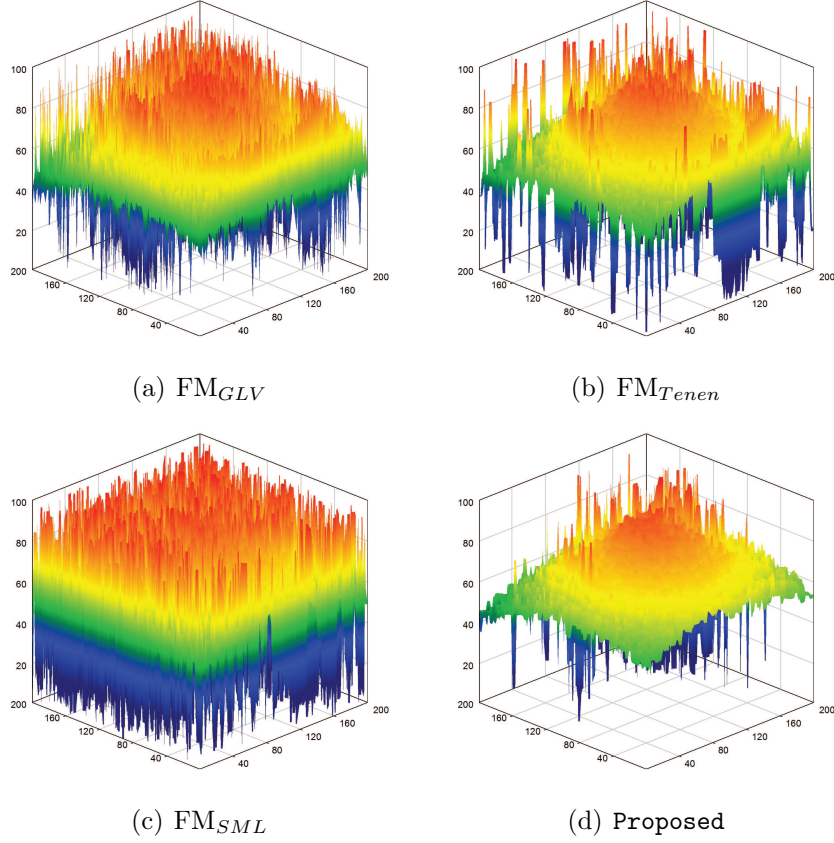


Figure 5.8:  
Depth map of Real Cone using various methods in presence of Gaussian noise of variance 0.001

The second main contribution to CCD noise is shot or salt and pepper noise. This noise generally introduce because of faulty switching during imaging as a sudden change of pixel value. To test the performance of the proposed method against the impulse change, we add salt and pepper noise having noise density 0.001 to the planner dataset. Fig. 5.9 show the depth map obtained by our proposed method and three other  $FM_{GLV}$ ,  $FM_{Tenen}$ ,  $FM_{SML}$  methods in presence of shot noise of density 0.001. From figure is clear that the effect of sudden change of pixel value due to shot noise in proposed method is quite better than the other methods and gives the more accurate depth compare to other techniques.



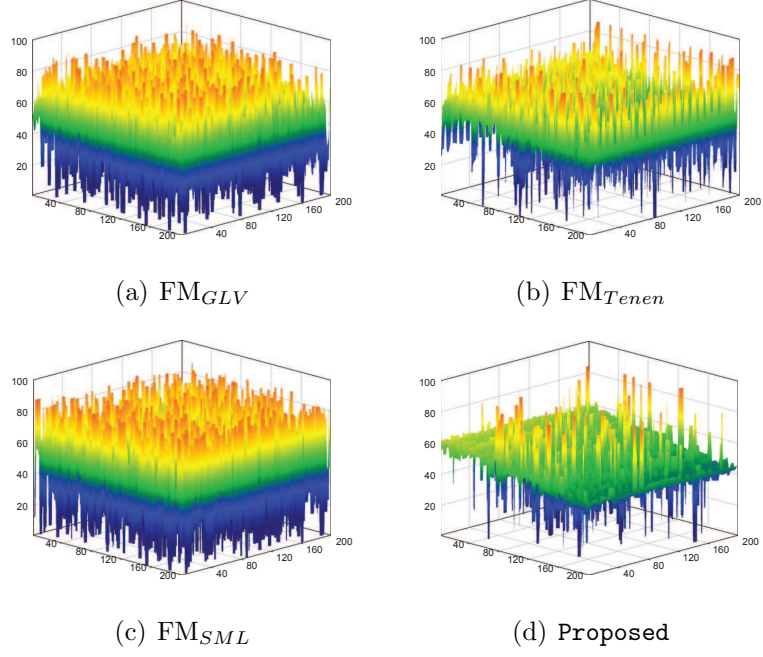


Figure 5.9:

Depth map of Planner Object using various methods in presence of shot noise of density 0.001

For the quantitative evaluation of the proposed method SSIM and  $Q_p$  criteria is used. Table 6.1 and 6.2 show the obtained results of the Simulated Cone, Real Cone and Planner Object in terms of SSIM and  $Q_p$  respectively. From results it is clear that the proposed method have more structural similarity as well as more edge information transferred from source images to a fused image, which indicates better fusion compare to other methods.

	$FM_{GLV}$	$FM_{Tenen}$	$FM_{SML}$	DT-CWT
Sim. Cone	0.1754	0.1870	0.1851	0.1881
Real Cone	0.4637	0.4536	0.4548	0.4734
Planner	0.4092	0.4275	0.4280	0.4399

Table 5.1: Structural similarity comparison

	$FM_{GLV}$	$FM_{Tenen}$	$FM_{SML}$	DT-CWT
Sim. Cone	0.1959	0.1807	0.2036	0.2134
Real Cone	0.2289	0.2516	0.2452	0.2642
Planner	0.1495	0.1754	0.1766	0.1895

Table 5.2: Edge information transformation comparison

The robustness of the proposed complex wavelet based focus measure system is tested by adding Gaussian noise of different variance. MSSIM and  $Q_p$  comparison at different Gaussian noise level of various focus measure operator for real cone and planner object sequence are shown in the fig. 5.10.

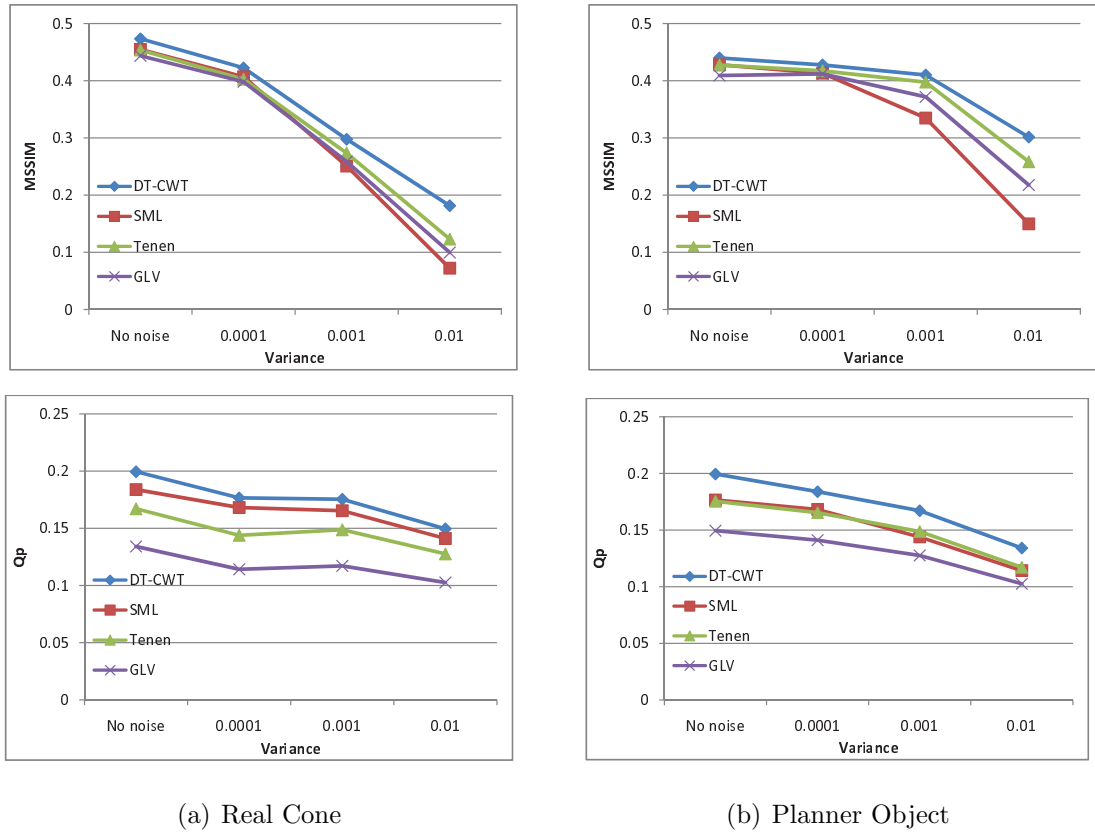


Figure 5.10: Comparison of focus measure (Gaussian noise)

As is seen in figure, for structural similarity measure  $FM_{SML}$  shows poor perfor-

mance while in case of edge information transformation  $FM_{GLV}$  method exhibits lower performance but, at each noise level the structural similarity and edge information transformation measure of the proposed method is higher. It is also seen that the rate of decrement of performance in proposed method is lower than the other well-known methods. It indicates that proposed method withstands more better against the noise compared to other methods.

## 5.4 Conclusion

An efficient depth map and 3D shape reconstruction technique based on dual tree complex wavelet transform is presented in this chapter. Shift-invariance nature and better directionality of DT-CWT helps to detect features precisely and distance transformation using exponentially decaying function gives more weight to the neighborhood pixels of the best candidate features. Experimental results in terms of MSSIM and  $Q_p$  validate the success of presented approach to estimate the depth of a scene compared to other well-documented methods in literature. Medical imaging, collision avoidance, shape reconstruction and object segmentation are some of the potential areas that can benefit from scheme presented in this chapter.

## Chapter 6

# Depth Map Estimation Using Dual-tree Complex Wavelets and a Quadrature Pair of Steerable Filters

The objective of shape from focus is to find out the depth of every point of the object from the camera lens. The depth of every point is calculated by finding the best focused points i.e. sharpest pixel, from a multi-focus image sequence. In this chapter a novel depth map estimation technique which is based on DT-CWT and a quadrature pair of steerable filter is presented. In proposed technique, by steering a quadrature pair of steerable filters, local oriented energy of the detected features of DT-CWT is measured, which helps to measure the focus pixel.

### 6.1 Quadrature Pair of Steerable Filter

Steerable filters [3] represent a class of filters which are synthesized as a linear combination of a set of basis filters. To analyze local orientation patterns and to measure the local energy along the dominant local direction in an image, a quadrature pair of steerable filters and the corresponding Hilbert transform pairs are used.

Any function  $g(x, y)$  which steers can be represented as linear sum of its own rotated versions, that is

$$g^\theta(x, y) = \sum_{m=1}^M k_m(\theta) g^{\theta_m}(x, y) \quad (6.1)$$

where  $k_m(\theta)$  refers to interpolation functions,  $g^{\theta_m}(x, y)$  forms basis filters and  $\theta$  is an arbitrary rotation.

The second derivative of Gaussian,  $G_2(x, y)$  [3], is used as a steerable filter in this paper. An arbitrary orientation of  $G_2(x, y)$  and its Hilbert transform  $H_2(x, y)$  are expressed as

$$\begin{aligned} G_2^\theta(x, y) &= k_1(\theta) G_2^0(x, y) + k_2(\theta) G_2^{\frac{\pi}{3}}(x, y) \\ &+ k_3(\theta) G_2^{\frac{2\pi}{3}}(x, y) \end{aligned} \quad (6.2)$$

$$\begin{aligned} H_2^\theta(x, y) &= l_1(\theta) H_2^0(x, y) + l_2(\theta) H_2^{\frac{\pi}{4}}(x, y) \\ &+ l_3(\theta) H_2^{\frac{\pi}{2}}(x, y) + l_4(\theta) H_2^{\frac{3\pi}{4}}(x, y) \end{aligned} \quad (6.3)$$

where normalized basis filters spaced equally between 0 and  $\pi$  with the following interpolation functions [3] for the Gaussian,  $k_{(\cdot)}(\theta)$ , and the Hilbert transform,  $l_{(\cdot)}(\theta)$ , respectively.

$$G_2^0(x, y) = 0.9213(2x^2 - 1)e^{-(x^2+y^2)} \quad (6.4)$$

$$G_2^{\frac{\pi}{3}}(x, y) = 1.843xye^{-(x^2+y^2)} \quad (6.5)$$

$$G_2^{\frac{2\pi}{3}}(x, y) = 0.9213(2y^2 - 1)e^{-(x^2+y^2)} \quad (6.6)$$

$$k_1(\theta) = (1 + 2 \cos 2\theta)/3 \quad (6.7)$$

$$k_2(\theta) = (1 + 2 \cos(2\theta - \frac{\pi}{3}))/3 \quad (6.8)$$

$$k_3(\theta) = (1 + 2 \cos(2\theta - \frac{4\pi}{3}))/3 \quad (6.9)$$

The quadrature filter of  $G_2$  is the Hilbert transform of it, which cannot steer. Therefore, a third-order polynomial is used for approximation and consequently the corresponding normalized basis filters of  $H_2(x, y)$  are given by

$$H_2^0(x, y) = (-2.205x + 0.978x^3)e^{-(x^2+y^2)} \quad (6.10)$$

$$H_2^{\frac{\pi}{4}}(x, y) = (-0.735y + 0.978x^2y)e^{-(x^2+y^2)} \quad (6.11)$$

$$H_2^{\frac{\pi}{2}}(x, y) = (-0.735x + 0.978xy^2)e^{-(x^2+y^2)} \quad (6.12)$$

$$H_2^{\frac{3\pi}{4}}(x, y) = (-2.205y + 0.978y^3)e^{-(x^2+y^2)} \quad (6.13)$$

$$l_1(\theta) = \cos^3 \theta \quad (6.14)$$

$$l_2(\theta) = -3 \cos^2 \theta \sin \theta \quad (6.15)$$

$$l_3(\theta) = -3 \cos \theta \sin^2 \theta \quad (6.16)$$

$$l_4(\theta) = -\sin^3 \theta. \quad (6.17)$$

## 6.2 Proposed Focus Measure Scheme

In SFF the success of any focus measure technique depends on its ability to calculate sharpness value of each image pixel. We take the advantage of dual-tree complex wavelet transform to decompose a pre-registered multi-focus image sequence to several directional subbands in complex wavelet domain. We then search for the best focus value using a quadrature pair of steerable filters. The decomposition in wavelet domain provides detailed information oriented in different directions. The intermediate subband ( $I$ ) with higher pixel value is computed employing the well known maximum selection rule on wavelet subbands of each image at the last level of the decomposition. After finding best feature information for each image local orientation energy ( $E$ ) for each pixel is measured. Local orientation energy is calculated by steering a quadrature pair of steerable filters and energy for each

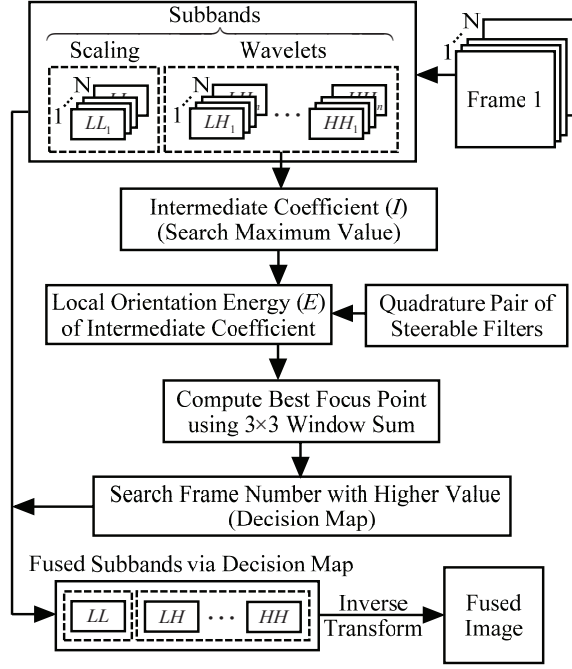


Figure 6.1: Block diagram of the proposed technique

pixel is given by

$$E(x, y, \theta) = [I(x, y) * G_2^\theta(x, y)]^2 + [I(x, y) * H_2^\theta(x, y)]^2 \quad (6.18)$$

Once the energy of each pixel is determined, using a small window around a pixel (x,y), focus measure (FM) is calculated and the value of pixel (x,y) is replaced by the sum of computed values to avoid abrupt fluctuations and measurement errors. That is

$$FM_W(i, j) = \sum_{x=i-N}^{i+N} \sum_{y=j-N}^{j+N} E(x, y, \theta) \quad (6.19)$$

where the subscript  $W$  refers to wavelet domain. Finally, decision map is constructed by comparing corresponding pixel values of each image; the frame with higher value of entire dataset is mapped onto the decision map. Using this decision map pixels are extracted from the scaling and wavelet coefficient of each image and then inverse transform yields the fused image. The entire process of the proposed technique for depth map estimation is depicted in Fig. 6.1.

### 6.3 Experiments and Results

In experiments, several multi-focus image sequence datasets such as simulated cone, real cone, planner object, pre-treatment, chess, clock, and book are used. The quantitative and visual results of the proposed technique are compared with two other well-known methods ( $FM_{SML}$  and  $FM_{GLV}$ ) as well as the method proposed in chapter 5. The sample frames of the datasets are shown in fig 4.4 and fig 5.3. The depth maps of the simulated cone and real cone data sets for various methods are depicted in fig. 6.2 and fig. 6.3. As seen in figures it is clear that the proposed steerable filter based focus measure technique gives uniform depth map without any spike compared to other methods. However, visually the depth maps for the proposed method and the method proposed in chapter 5 look similar, but the smoothness in the depth map of the proposed method is better.

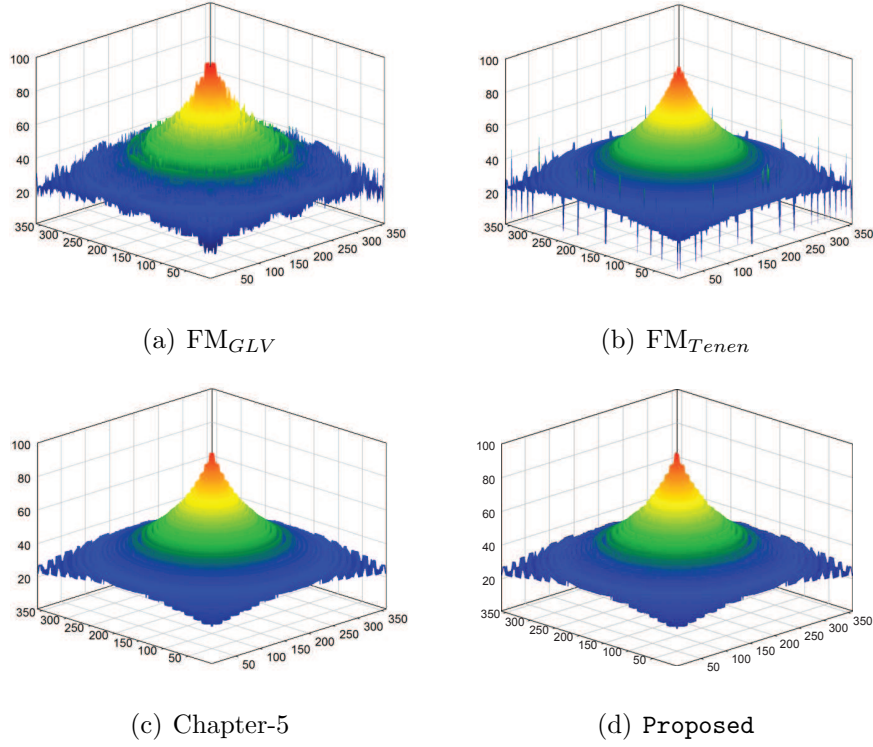


Figure 6.2: Depth map of Simulated Cone sequence using various methods



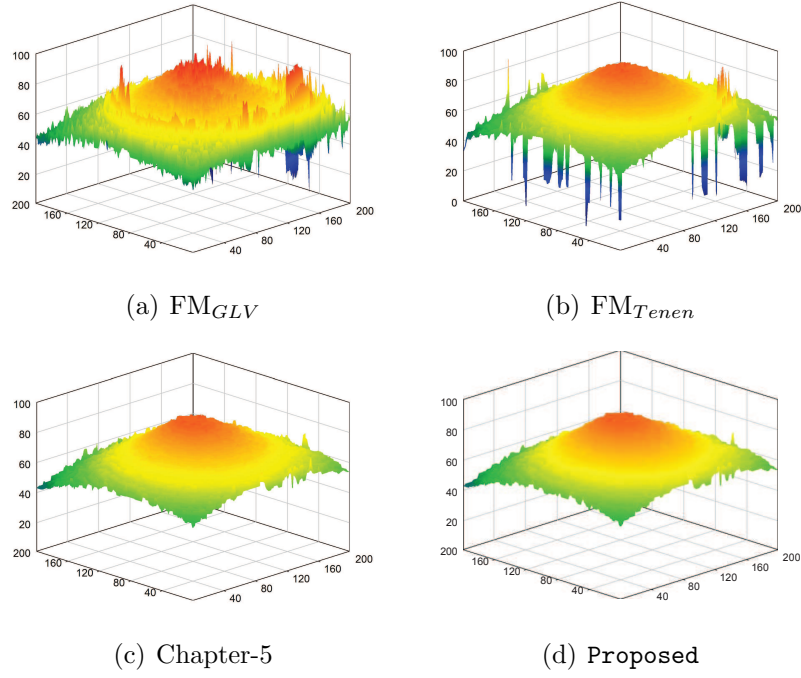


Figure 6.3: Depth map of Real Cone dataset using various methods

For quantitative evaluation of proposed method, structural similarity and edge information transformation measure is used. Proposed method is tested with simulated cone, real cone, planner object, pre-treatment, chess, clock and book datasets and the results of SSIM and  $Q_p$  are compared with three focus measure methods including methods proposed in chapter 5 and tabulated in table 6.1 and 6.2.

From tables it is seen that, proposed has more structural similarity and more edge information transformation between input images and a fused image. However, it is also seen that the improvement of SSIM and  $Q_p$  measure between proposed and method proposed in chapter 5 for simulated cone, real cone and planner datasets are small, but for pre-treatment, clock and book datasets, performance improvement is higher. One of the possible reason for this higher improvement is that in these datasets images have higher flat area. Overall performance improvement of the proposed is better than the other well-documented methods.

	$FM_{GLV}$	$FM_{SML}$	Chapter-5	Propoed
Sim. Cone	0.1754	0.1851	0.1881	0.1916
Real Cone	0.4637	0.4548	0.4734	0.4812
Planner	0.4092	0.4280	0.4399	0.4425
Pre-treat	0.5414	0.5455	0.5829	0.5889
Chess	0.7111	0.7082	0.7812	0.7916
Clock	0.8630	0.8797	0.8812	0.8873
Book	0.8662	0.8728	0.8796	0.8973

Table 6.1: Structural similarity comparison

	$FM_{GLV}$	$FM_{SML}$	Chapter-5	Proposed
Sim. Cone	0.1959	0.2036	0.2134	0.2218
Real Cone	0.2289	0.2452	0.2642	0.2726
Planner	0.1495	0.1766	0.1895	0.1927
Pre-treat	0.3473	0.3875	0.4112	0.4269
Chess	0.3172	0.2449	0.3196	0.3234
Clock	0.7033	0.7667	0.7832	0.7922
Book	0.6788	0.7487	0.7538	0.7642

Table 6.2: Edge information transformation comparison

## 6.4 Conclusion

In this chapter, we have introduced an efficient focus measure technique for depth map estimation based on the use of complex wavelet subbands to extract detailed feature information from a multi-focus image sequence. Local oriented energy of the detected features is calculated using quadrature pair of steerable filters. Steerable filters remove inherent limitations of traditional gradient detection based techniques which perform inadequately for oriented intensity varying scenarios. Simulation results show the performance improvement of DT-CWT based technique employing steerable filters in terms of fusion assessment factors and visual perception. The feature extraction and exploiting the best candidate pixel fits the use of dual-tree complex wavelet transform. Shown in the contour plane, the boundary of depth are accurately extracted compare with previous techniques. This is due to better directional-selectivity and shift-invariance attribute of DT-CWT in general, and full symmetry of the  $9/7-10/8$  filter bank with promising number of vanishing moments, in particular. In view of success of the presented approach in depth map estimation, it is reasonable to hope promising results, employing the proposed framework, in relevant applications such as 3D shape reconstruction and object identification.

## Chapter 7

# Conclusion and Future Works

### 7.1 Conclusion

This chapter summarises and concludes the investigation of depth map estimation by image focus presented in this thesis. General recommendations for further work and development in the area of depth map estimation using image focus are proposed in the second section of this chapter.

In this thesis depth map estimation based on features extraction is investigated in terms of development of novel multi-focus image fusion algorithms. A broad overview of the developments of this type, presented in technical literature so far, is given in chapter 2. The quantitative fusion assessment techniques is explained in chapter 3. In chapter 4 depth map estimation based on SUSAN operator and an exponentially decaying function is proposed. The excellent capacity of feature detection and structure preserving noise filtering using SUSAN, identifies features precisely while exponentially decaying function gives more weight to the neighborhood pixel of the detected features, which helps to measure the focus part of the image. Further more taking the advantage of DT-CWT in terms of shift-invariance and better directionality, other two efficient depth map estimation techniques using distance transformation and a quadrature pair of steerable filters are presented in chapter 5 and 6 respectively. As better directionality attribute,

DT-CWT is able to detect features more efficiently. An exponentially decaying function gives more weight to the near vicinity of the detected feature pixels while a quadrature pair of steerable filters measures the local orientation energy of the detected features to calculate the focus pixels.

The depth map as well as fused imagery obtained using these novel approaches are evaluated against conventional depth map estimation methods, in subjective as well as objective preference tests which show a clear advantage of the proposed method.

## 7.2 Future Works

The investigation into the field of depth map estimation based on feature detection is presented in this thesis, however, it was not exhaustive and some natural extensions to the research presented in it are recommended in this section.

For more practical depth map estimation research, large number of datasets would be required. These datasets would ideally include imagery of as many possible scenes from wide range of application scenarios.

Presented feature based approaches of depth map estimation required more effort to deal with the flat regions. As, flat regions are the regions without features, there may possible that these feature based approaches of depth map estimation not work efficiently. Also, more attention is required towards the robustness of the focus measure system in terms of dealing with different kind and level of noises.

Finally, construction of hardware platforms for real time depth map estimation implementation, where fusion algorithms can be tested in real life conditions, is highly recommended.

An applications such as robot guidance, collision avoidance, 3D feature extraction, medical imaging, range segmentation and microscopic imaging would bring benefits to the proposed focus measure systems.

## References

- [1] M.B. Ahmad and T.S. Choi. A heuristic approach for finding best focused shape. *IEEE Trans. Circuits Syst. Video Technol.*, 15(4):566–574, 2005.
- [2] M. Asif and T.S. Choi. Shape from focus using multilayer feedforward neural network. *IEEE Trans. Image Processing*, 10(11):1670–1675, 2001.
- [3] W. T. Freeman and E. H. Adelson. The design and use of steerable filters. *IEEE Trans. on PAMI*, 13(9):891–906, 1991.
- [4] D. Zhang G.H. Qu and P. Yan. Information measure for performance of image fusion. *Electronics Letters*, 38(7):313–315, 2002.
- [5] B.S. Manjunath H. Li and S.K. Mitra. Multisensor image fusion using the wavelet transform. *In Proc. of the IEEE ICIP*, 1:51–55, 1994.
- [6] F.S. Helmli and S. Scherer. Adaptive shape from focus with an error estimation in light microscopy. *International Symposium on ISPA*, pages 188–193, 2001.
- [7] W. Huang and Z. Jing. Multi-focus image fusion using pulse coupled neural network. *Pattern Recognition Letters*, 28:1123–1132, 2007.
- [8] N. G. Kingsbury. Complex wavelets for shift invariant analysis and filtering of signals. *Journal of Applied and Computational Harmonic Analysis*, 10(3):234–253, 2001.

- [9] E. Krotov. Focusing. *International Journal of Computer Vision*, 1:223–237, 1987.
- [10] M.T. Mahmood and T.S. Choi. A feature analysis approach to estimate 3d shape from image focus. *IEEE ICIP*, pages 3216–3219.
- [11] A.S. Malik and T.S. Choi. A novel algorithm for estimation of depth map using image focus for 3d shape recovery in the presence of noise. *Pattern Recognition*, 41(7).
- [12] A.S. Malik and T.S. Choi. Finding best focused points using intersection of two lines. *International Conference on Image Processing*, pages 1952–1955, 2008.
- [13] S. Shim M.T. Mehmood and T.S. Choi. Wavelet and pca based approach for 3d shape recovery from image focus. *Proceedings of SPIE*, pages 1496–1499, 2008.
- [14] S.K. Nayyar and Y. Nakagawa. Shape from Focus: an effective approach for rough surfaces. *International Conference on Robotics and Automation*, 2:218–225, 1990.
- [15] S.K. Nayyar and Y. Nakagawa. Shape from focus. *IEEE Trans. on PAMI*, 16(8):824–831, 1994.
- [16] K.S. Pradeep and A.N. Rajangopalan. Improving shape from focus using defocus cues. *IEEE Trans. on Image Processing*, 16(7):1920–1925, 2007.
- [17] I. W. Selesnick. The design of approximate hilbert transform pairs of wavelet bases. *IEEE Trans. on Signal Process*, 50(5):1144–1152, 2002.
- [18] S.M. Smith and J.M. Brady. SUSAN - A new approach to low level image processing. *International journal of computer vision*, 23(1):45–78, 1997.

- [19] M. Subbarao and T.S. Choi. Accurate recovery of three dimensional shape from image focus. *IEEE Trans. Pattern Anal. Mach. Intell.*, 17(3):266–274, 1995.
- [20] M. Asif T. Choi and J. Yun. Three-dimensional shape recovery from focused image surface. *IEEE ICASSP*, 6:3269–3272, 1999.
- [21] S.K. Rogers T.A. Wilson and M. Kabrisky. Perceptual-based image fusion for hyperspectral data. *IEEE Trans. on Geoscience and Remote Sensing*, 35(4).
- [22] D. B. H. Tay. ETHFB: A new class of even-length biorthogonal wavelet filters for hilbert pair design. *IEEE Trans. on Circ. Syst. I*, 50(6):1580–1588, 2008.
- [23] J.M. Tenenbaum. Accomodations in computer vision. *PhD Thesis, Stanford University*, 1970.
- [24] C.S. Xydeas and V. Petrovic. Objective image fusion performance measure. *Electronics Letters*, 36(4):308–309, 2000.
- [25] R. Yu and A. Baradarani. Sampled-data design of FIR dual filter banks for dual-tree complex wavelet transforms via LMI optimization. *IEEE Trans. on Signal Processing*, 56(7–2):3369–33752, 2008.
- [26] J. Yun and T.S. Choi. Accurate 3-d shape recovery using curved window focus measure. *IEEE Conference on ICIP*, pages 910–914, 1999.
- [27] H.R. Sheikh Z. Wang, A.C. Bovik and E.P. Simoncelli. Image quality assessment: From error measurement to structural similarity. *IEEE Trans. on Image Processing*, 13(1):600–612, 2004.
- [28] S.H. Sun Z.L. Zhang and F.C. Zheng. Image fusion based on median filters and SOFM neural networks: A three-step scheme. *Signal Processing*, 81(6):1325–1330, 2001.



# Vita Auctoris

**NAME** : Pankajkumar Mendapara

**BIRTH DATE** : 19<sup>th</sup> Nov. 1976

**BIRTH PLACE** : INDIA

**EDUCATION**

**2008–2010** : **Master of Applied Science**  
Electrical and Computer Engineering  
University of Windsor, Windsor, ON Canada

**1999–2003** : **Bachelor of Engineering**  
Saurashtra University  
Rajkot (Gujarat) India

**1997–1999** : **Diploma Electrical Engineering**  
Government Polytechnique  
Rajkot (Gujarat) India

**1994–1997** : **Bachelor of Science**  
Sardar Patel University  
V.V.Nagar (Gujarat) India

**WORK EXPERIENCE**

**2003–2008** : **Research and Development Engineer**  
Setu Industries  
Rajkot (Gujarat) India